

# Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/JP2006/302092

International filing date: 07 February 2006 (07.02.2006)

Document type: Certified copy of priority document

Document details: Country/Office: JP  
Number: 2005-069768  
Filing date: 11 March 2005 (11.03.2005)

Date of receipt at the International Bureau: 22 March 2006 (22.03.2006)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日  
Date of Application: 2005年 3月11日

出 願 番 号  
Application Number: 特願2005-069768

パリ条約による外国への出願  
に用いる優先権の主張の基礎  
となる出願の国コードと出願  
番号

The country code and number  
of your priority application,  
to be used for filing abroad  
under the Paris Convention, is

J P 2005-069768

出 願 人  
Applicant(s): 日本電信電話株式会社

2006年 3月 8日

特許庁長官  
Commissioner,  
Japan Patent Office

中 嶋



【書類名】 特許願  
【整理番号】 NTTH166666  
【提出日】 平成17年 3月11日  
【あて先】 特許庁長官殿  
【国際特許分類】 G10L  
【発明者】  
    【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内  
    【氏名】 澤田 宏  
【発明者】  
    【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内  
    【氏名】 荒木 章子  
【発明者】  
    【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内  
    【氏名】 向井 良  
【発明者】  
    【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内  
    【氏名】 牧野 昭二  
【特許出願人】  
    【識別番号】 000004226  
    【氏名又は名称】 日本電信電話株式会社  
【代理人】  
    【識別番号】 100066153  
    【弁理士】  
    【氏名又は名称】 草野 卓  
【選任した代理人】  
    【識別番号】 100100642  
    【弁理士】  
    【氏名又は名称】 稲垣 稔  
【手数料の表示】  
    【予納台帳番号】 002897  
    【納付金額】 16,000円  
【提出物件の目録】  
    【物件名】 特許請求の範囲 1  
    【物件名】 明細書 1  
    【物件名】 図面 1  
    【物件名】 要約書 1  
    【包括委任状番号】 9806848

【書類名】 特許請求の範囲

【請求項 1】

複数の信号源から発せられた信号が混合された混合信号から所定の信号を抽出する信号抽出装置であって、

前記混合信号を時間周波数領域の混合信号に変換する時間周波数領域変換手段と、

前記時間周波数領域の混合信号を用い、周波数毎に分離行列と分離信号とを算出する信号分離手段と、

前記分離行列の一般化逆行列の列である基底ベクトルを正規化し、その正規化された基底ベクトルをクラスタリングし、そのクラスタの分散を指標として前記分離信号から目的信号を選択する目的信号選択手段と、

を有することを特徴とする信号抽出装置。

【請求項 2】

請求項 1 に記載の信号抽出装置であって、

更に、前記時間周波数領域の混合信号と前記基底ベクトルとを用いて時間周波数マスクを生成するマスク生成手段と、

前記時間周波数マスクを前記目的信号選択手段で選択された前記目的信号に適用し、マスク処理目的信号を生成するマスキング手段と、

を有することを特徴とする信号抽出装置。

【請求項 3】

請求項 2 に記載の信号抽出装置であって、

前記マスク生成手段が、

前記時間周波数領域の混合信号を用いて白色化行列を生成する白色化行列生成手段と、

前記白色化行列を用い、前記時間周波数領域の混合信号を要素とする混合信号ベクトルを変換した白色化混合信号ベクトルと、前記基底ベクトルを変換した白色化基底ベクトルとを算出する白色化手段と、

前記白色化混合信号ベクトルと前記白色化基底ベクトルとがなす角度を時間周波数毎に算出する角度計算手段と、

前記角度を要素とする関数である前記時間周波数マスクを生成する関数演算手段と、

を有することを特徴とする信号抽出装置。

【請求項 4】

請求項 3 に記載の信号抽出装置であって、

前記白色化行列が、

周波数を  $l$  とし、離散時間を  $\tau$  とし、前記混合信号ベクトルを  $X(l, \tau)$  とし、ベクトル  $*$  を時間平均したベクトルを  $\langle * \rangle_\tau$  とし、ベクトル  $*$  の複素共役転置ベクトルを  $*^H$  とし、 $R(l) = \langle X(l, \tau) \cdot X(l, \tau)^H \rangle_\tau$  とした場合における  $V(l) = R(l)^{-1/2}$  であり、

前記白色化手段が、

$Z(l, \tau) = V(l) \cdot X(l, \tau)$  の演算により、前記白色化混合信号ベクトル  $Z(l, \tau)$  を算出し、前記基底ベクトルを  $A(l)$  とした場合における  $B(l) = V(l) \cdot A(l)$  の演算により、前記白色化基底ベクトル  $B(l)$  を算出する手段であり、

前記角度計算手段が、

ベクトル  $*$  の絶対値を  $|*|$  とし、ベクトル  $*$  のノルムを  $\|*\|$  とした場合における  $\theta(l, \tau) = \cos^{-1}(|B(l)^H \cdot Z(l, \tau)| / \|B(l)\| \cdot \|Z(l, \tau)\|)$  の演算により、前記角度  $\theta(l, \tau)$  を算出する手段であり、

前記関数演算手段が、

$\alpha, \beta, \theta_T$  を実数パラメータとした場合におけるロジスティック関数  $M(\theta(l, \tau)) = \alpha / (1 + e^{\beta \cdot (\theta(l, \tau) - \theta_T)})$  を前記時間周波数マスクとして算出する手段である、

ことを特徴とする信号抽出装置。

【請求項 5】

請求項 1 に記載の信号抽出装置であって、

前記目的信号選択手段が、

前記基底ベクトル  $A_p(l)$  ( $p = 1, \dots, N$ 、 $N$ は前記信号源の数)の各要素  $A_{qp}(l)$  ( $q = 1, \dots, M$ 、 $M$ は前記混合信号を観測するセンサの数)に対し、 $\exp$ をネピア数とし、 $\arg$ を偏角とし、 $f$ を周波数とし、 $j$ を虚数単位とし、 $c$ を信号の伝達速度とし、 $Q$ を  $M$ 以下の自然数から選択された基準値とし、 $d$ を実数パラメータとした場合における

【数 1】

$$A_{qp}'(f) = |A_{qp}(f)| \exp \left[ j \frac{\arg[A_{qp}(f)/A_{Qp}(f)]}{4fc^{-1}d} \right]$$

の演算を行う周波数正規化手段を有する、  
ことを特徴とする信号抽出装置。

【請求項 6】

請求項 5 に記載の信号抽出装置であって、  
前記実数パラメータ  $d$  が、  
前記基準値  $Q$  に対応する基準センサと他のセンサとの最大距離  $d_{\max}$  である、  
ことを特徴とする信号抽出装置。

【請求項 7】

請求項 1 に記載の信号抽出装置であって、  
前記目的信号選択手段が、  
分散が最小となるクラスタを選択し、選択されたクラスタに対応する前記分離信号を目的信号として選択する手段である、  
ことを特徴とする信号抽出装置。

【請求項 8】

複数の信号源から発せられた信号が混合された混合信号から所定の信号を抽出する信号抽出方法であって、  
時間周波数領域変換手段が、入力された前記混合信号を時間周波数領域の混合信号に変換して出力する手順と、  
信号分離手段が、入力された前記時間周波数領域の混合信号を用い、周波数毎に分離行列と分離信号とを算出して出力する手順と、  
目的信号選択手段が、入力された前記分離行列の一般化逆行列の列である基底ベクトルを正規化し、その正規化された基底ベクトルをクラスタリングし、そのクラスタの分散を指標として、入力された前記分離信号から目的信号を選択する手順と、  
を有することを特徴とする信号抽出方法。

【請求項 9】

請求項 1 から 7 の何れかに記載の信号抽出装置としてコンピュータを実行させるための信号抽出プログラム。

【請求項 10】

請求項 9 記載の信号抽出プログラムを格納したコンピュータ読み取り可能な記録媒体。

【書類名】 明細書

【発明の名称】 信号抽出装置、信号抽出方法、信号抽出プログラム及び記録媒体

【技術分野】

【0001】

本発明は、信号処理の技術分野に属し、特に、複数の信号が空間内で混合されたものから目的の信号のみを抽出し、その他の信号を抑圧する技術に関する。

【背景技術】

【0002】

複数のセンサを用いることによりある信号を抽出してその他の信号を抑圧する従来技術として、ビームフォーマ（beamformer）（ビームフォーミング（beamforming）とも呼ぶ）が広く知られている（例えば、非特許文献1参照）。しかし、ビームフォーマでは、目的信号の方向に関する情報を必要とするため、そのような情報が得られない（または推定できない）状況では利用しにくいという欠点がある。

また、より新しい別の技術としては、ブラインド信号分離（BSS：Blind Signal Separation）が挙げられる（例えば、非特許文献2参照）。これは、上記ビームフォーマが必要とした情報を必要としないという点で優れており、様々な状況での利用が期待されている。

【非特許文献1】 B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to special filtering," IEEE ASSP Magazine, pp. 2-24, April 1988

【非特許文献2】 S. Araki, S. Makino, A. Blin, R. Mukai, and H. Sawada, "Undetermined blind separation for speech in real environments with sparseness and ICA," in Proc. ICASSP 2004, vol. III, May 2004, pp. 881-884

【発明の開示】

【発明が解決しようとする課題】

【0003】

しかし、ブラインド信号分離では、センサでの観測信号を複数の信号に分離して出力する。そのうちの一部分が目的とした信号であるが、ブラインド信号分離では、どの信号が目的信号であるかという情報までは提供されない。従って、何らかの別の手段によって、どれが目的信号であるかを判定しなければならない。

本発明はこのような点に鑑みてなされたものであり、目的信号の方向に関する情報を必要とせずに、複数の信号源から発せられた信号が混合された混合信号から目的信号を抽出する技術を提供することを目的とする。

【課題を解決するための手段】

【0004】

本発明では上記課題を解決するために、まず、複数の信号源から発せられた信号が混合された混合信号を時間周波数領域の混合信号に変換し、時間周波数領域の混合信号を用い、周波数毎に分離行列と分離信号とを算出する。次に、分離行列の一般化逆行列の列である基底ベクトルを正規化し、その正規化された基底ベクトルをクラスタリングし、そのクラスタの分散を指標として分離信号から目的信号を選択する。なお、分離行列が正方行列の場合、「一般化逆行列」は逆行列に相当する。すなわち、ここでの一般化逆行列は、通常の逆行列をも含む概念である。

【0005】

ここで、基底ベクトルの正規化は、複数の信号源から発せられた信号の畳み込み混合を所定のモデル（例えば、近距離場モデル）に近似した場合に、正規化された基底ベクトルが信号源の位置のみに依存したクラスタを形成するように行われる。しかし、実環境では、このようなモデルに反映されない様々な因子が存在する。例えば、近距離場モデルでは、信号が壁などに反射することによって生じる伝送歪みが反映されない。このような実環境とモデルとの乖離は、信号源からセンサまでの距離が大きいほど大きくなり、センサに近い信号ほど小さくなる。従って、センサに近い信号ほど実環境に近い条件で正規化ができ、実環境とモデルとの乖離に起因するクラスタの分散を小さくできる。本発明では、こ

の関係に着目し、クラスタの分散の大きさを指標として、センサに近い信号を目的信号として抽出する。

#### 【0006】

以上の操作により、目的信号を抽出し、他の妨害信号をある程度抑圧することができる。しかし、分離行列及び分離信号の算出を独立成分分析(ICA: Independent Component Analysis)で行う場合、以上の処理によって完全に抑制できる妨害信号の数はセンサの数から1を引いた数までである。すなわち、妨害信号の数がそれ以上の場合は抑圧しきれない残留成分が残る。そのため、本発明において好ましくは、更に、時間周波数領域の混合信号と基底ベクトルとを用いて時間周波数マスクを生成し、この時間周波数マスクを選択された目的信号に適用し、マスク処理目的信号を生成する。これにより信号源の数がセンサ数より多い場合でも妨害信号をより良く抑圧することができる。

#### 【0007】

更に、本発明において好ましくは、時間周波数領域の混合信号を用いて白色化行列を生成し、白色化行列を用い、時間周波数領域の混合信号を要素とする混合信号ベクトルを変換した白色化混合信号ベクトルと、基底ベクトルを変換した白色化基底ベクトルとを算出し、白色化混合信号ベクトルと白色化基底ベクトルとがなす角度を時間周波数毎に算出し、この角度を要素とする関数を時間周波数マスクとして生成する。この時間周波数マスクを目的信号に適用することにより、目的信号が有する妨害信号を更に抑圧できる。

#### 【0008】

また更に、本発明において好ましくは、白色化行列が、周波数を $l$ とし、離散時間を $\tau$ とし、混合信号ベクトルを $X(l, \tau)$ とし、ベクトル $*$ を時間平均したベクトルを $\langle * \rangle_\tau$ とし、ベクトル $*$ の複素共役転置ベクトル(各要素の複素共役を採り転置したベクトル)を $*^H$ とし、 $R(l) = \langle X(l, \tau) \cdot X(l, \tau)^H \rangle_\tau$ とした場合における $V(l) = R(l)^{-1/2}$ である。そして、 $Z(l, \tau) = V(l) \cdot X(l, \tau)$ の演算により、白色化混合信号ベクトル $Z(l, \tau)$ を算出し、基底ベクトルを $A(l)$ とした場合における $B(l) = V(l) \cdot A(l)$ の演算により、白色化基底ベクトル $B(l)$ を算出する。また、ベクトル $*$ の絶対値を $|*|$ とし、ベクトル $*$ のノルムを $\|* \|$ とした場合における $\theta(l, \tau) = \cos^{-1}(|B(l)^H \cdot Z(l, \tau)| / \|B(l)\| \cdot \|Z(l, \tau)\|)$ の演算により、角度 $\theta(l, \tau)$ を算出し、 $\alpha, \beta, \theta_T$ を実数パラメータとした場合におけるロジスティック関数 $M(\theta(l, \tau)) = \alpha / (1 + e^{\beta \cdot (\theta(l, \tau) - \theta_T)})$ を時間周波数マスクとして算出する。この時間周波数マスクを抽出した目的信号に適用することにより、目的信号が有する妨害信号を更に抑圧できる。

#### 【0009】

また、本形態において好ましくは、基底ベクトル $A_p(l)$  ( $p = 1, \dots, N$ 、 $N$ は信号源の数)の各要素 $A_{qp}(l)$  ( $q = 1, \dots, M$ 、 $M$ は混合信号を観測するセンサの数)に対し、 $\exp$ をネピア数とし、 $\arg$ を偏角とし、 $f$ を周波数とし、 $j$ を虚数単位とし、 $c$ を信号の伝達速度とし、 $Q$ を $M$ 以下の自然数から選択された基準値とし、 $d$ を実数パラメータとした場合における、

#### 【数2】

$$A_{qp}'(f) = |A_{qp}(f)| \exp \left[ j \frac{\arg[A_{qp}(f)/A_{Qp}(f)]}{4fc^{-1}d} \right] \dots (1)$$

の演算によって基底ベクトルの周波数正規化を行う。この正規化により、複数の信号源から発せられた信号の畳み込み混合を所定のモデルに近似した場合に、正規化された基底ベクトルが信号源の位置に依存したクラスタを形成することとなる。その結果、上述のようにクラスタの分散の大きさを指標として目的信号を抽出することが可能となる。

#### 【0010】

また更に、本発明において好ましくは、上述の式(1)における実数パラメータ $d$ は、前記基準値 $Q$ に対応する基準センサと他のセンサとの最大距離 $d_{\max}$ である。これにより、周波数正規化ベクトルの各要素 $A_{qp}'(f), A_{q'p'}'(f)$ 間における $|\arg(A_{qp}'(f)) - \arg(A_{q'p'}'(f))|$ が

$g(A_{q,p}^{(i')})$ の増加に対して $|A_{q,p}^{(i)} - A_{q,p}^{(i')}|$ も単調増加するという関係を維持しつつ、できるだけ広い範囲にクラスタを分散させることができる（詳細は後述）。その結果、クラスタリングの精度を向上させ、より正確に目的信号を抽出することが可能となる。

また、本発明において好ましくは、分散が最小となるクラスタを選択し、選択されたクラスタに対応する分離信号を目的信号として選択する。これにより、センサに最も近い信号を目的信号として抽出することができる。

#### 【発明の効果】

##### 【0011】

本発明により、目的信号の方向に関する情報を必要とせずに、複数の信号源から発せられた信号が混合された混合信号から目的信号を抽出することが可能となる。

#### 【発明を実施するための最良の形態】

##### 【0012】

以下、本発明の実施の形態を図面を参照して説明する。

##### （第1の実施の形態）

まず、本発明における第1の実施の形態について説明する。

##### <ハードウェア構成>

図1は、本形態における信号抽出装置1のハードウェア構成を例示したブロック図である。

図1に例示するように、この例の信号抽出装置1は、CPU（Central Processing Unit）10、入力部20、出力部30、補助記憶装置40、RAM（Random Access Memory）50、ROM（Read Only Memory）60及びバス70を有している。

##### 【0013】

この例のCPU10は、制御部11、演算部12及びレジスタ13を有し、レジスタ13に読み込まれた各種プログラムに従って様々な演算処理を実行する。また、この例の入力部20は、データが入力される入力ポート、キーボード、マウス等であり、出力部30は、データを出力する出力ポート、ディスプレイ等である。補助記憶装置40は、例えば、ハードディスク、MO（Magneto-Optical disc）、半導体メモリ等であり、本形態の信号分離処理を実行するための信号抽出プログラムを格納した信号抽出プログラム領域41及びセンサで観測された時間領域の混合信号等の各種データが格納されるデータ領域42を有している。また、RAM50は、例えば、SRAM（Static Random Access Memory）、DRAM（Dynamic Random Access Memory）等であり、信号抽出プログラムが書き込まれる信号抽出プログラム領域51及び各種データが書き込まれるデータ領域52を有している。また、この例のバス70は、CPU10、入力部20、出力部30、補助記憶装置40、RAM50及びROM60を通信可能に接続している。

##### 【0014】

##### <ハードウェアとソフトウェアとの協働>

この例のCPU10は、読み込まれたOS（Operating System）プログラムに従い、補助記憶装置40の信号抽出プログラム領域41に格納されている信号抽出プログラムを、RAM50の信号抽出プログラム領域51に書き込む。同様にCPU10は、補助記憶装置40のデータ領域42に格納されている時間領域の混合信号等の各種データをRAM50のデータ領域52に書き込む。さらに、CPU10は、この信号抽出プログラムや各種データが書き込まれたRAM50上のアドレスをレジスタ13に格納する。そして、CPU10の制御部11は、レジスタ13に格納されたこれらのアドレスを順次読み出し、読み出したアドレスが示すRAM50上の領域からプログラムやデータを読み出し、そのプログラムが示す演算を演算部12に順次実行させ、その演算結果をレジスタ13に格納していく。

##### 【0015】

図2は、このようにCPU10に信号抽出プログラムが読み込まれることにより構成される信号抽出装置1のブロック図の例示である。



この図に例示するように、信号抽出装置1は、記憶領域101～114を有するメモリ100、時間周波数領域変換部120、信号分離部130、目的信号選択部140、時間周波数マスキング部150、時間領域変換部160、制御部170及び一時メモリ180を有する。ここでメモリ100及び一時メモリ180は、例えば、レジスタ13、補助記憶装置40のデータ領域42或いはRAM50のデータ領域52の少なくとも一部に相当するが、特にこれに限定されない。また、時間周波数領域変換部120、信号分離部130、目的信号選択部140、時間周波数マスキング部150、時間領域変換部160、制御部170及び一時メモリ180は、例えば、CPU10にOSプログラムや信号抽出プログラムが読み込まれることにより構成されるものである。

#### 【0016】

図3(a)は、図2における目的信号選択部140の詳細構成を例示したブロック図であり、図3(b)は、図3(a)における基底ベクトルクラスタリング部142の詳細構成を例示したブロック図である。

これらの図に例示するように、目的信号選択部140は、逆行列算出部141、基底ベクトルクラスタリング部142及び選択部143を有している。また、基底ベクトルクラスタリング部142は、周波数正規化部142a、ノルム正規化部142b、クラスタリング部142c及び分散判定部142dを有している。

#### 【0017】

図4(a)は、図2における時間周波数マスキング部150の詳細構成を例示したブロック図であり、図4(b)は、図4(a)におけるマスク生成部151の詳細を例示したブロック図である。

これらの図に例示するように、時間周波数マスキング部150は、マスク生成部151及びマスキング部152を有している。また、マスク生成部151は、白色化行列生成部151a、白色化部151b、角度計算部151c及び関数演算部151dを有している。

#### 【0018】

なお、図2～図4における実線の矢印は実際のデータの流れを示し、破線の矢印は理論上の情報の流れを示すが、制御部170及び一時メモリ180に出入りするデータの流れは省略してある。また、信号抽出装置1は制御部170の制御のもと各処理を実行する。また、特に示さない限り、制御部170は、必要なデータを一時メモリ180に読み書きしながら各処理を実行していく。

#### 【0019】

##### <処理>

次に、本形態における信号抽出装置1の処理について説明する。

##### 〔センサでの観測信号（時間領域の混合信号）〕

まず、センサで信号がどのように観測されるかを説明する。

空間内にN個の信号源 $k$  ( $k \in \{1, 2, \dots, N\}$ )が存在し、それらの信号 $s_k(t)$  ( $t$ はサンプリング時刻)が混合され、M個のセンサ $q$  ( $q \in \{1, 2, \dots, M\}$ )で混合信号 $x_q$ として観測されたとする。なおここでは、信号源 $k$ からセンサ $q$ までの距離により信号が減衰・遅延し、また壁などにより信号が反射して伝送歪みが発生しうる状況を扱う。このような状況での混合は、信号源 $k$ からセンサ $q$ へのインパルス応答 $h_{qk}(r)$ による畳み込み混合

##### 【数3】

$$x_q(t) = \sum_{k=1}^N \sum_{r=0}^{\infty} h_{qk}(r) s_k(t-r) \quad \cdots (2)$$

となる。ここで、 $r$ は掃引のための変数を示している。

#### 【0020】

本形態では、観測信号 $x_1(t), \dots, x_M(t)$ のみから、何れかの信号源から発せられた目的信号を抽出してその他の妨害信号を抑圧した信号 $y(t)$ を算出する。なお、信号源の数 $N$ はセンサの数 $M$ よりも多くても少なくてもよく、また同数であってもよい。さらに、信号源

の数Nを事前に知る必要もなく、また信号源の数が数えられないような状況であっても構わない。

【0021】

〔時間周波数領域〕

本形態では、このような畳み込み混合でモデル化される混合信号から目的信号を抽出するわけであるが、この畳み込み混合の問題は扱いが繁雑である。そのため、短時間フーリエ変換（STFT: Short-Time Fourier Transform）を用い、時間領域の混合信号を取扱いが容易な時間周波数領域の混合信号に変換してから各処理を実行する。次に、この時間周波数領域の混合信号をモデル化する。

時間周波数領域では、式（2）で示される時間領域での畳み込み混合が、

【数4】

$$X_q(f, \tau) = \sum_{k=1}^N H_{qk}(f) S_k(f, \tau) \quad \cdots (3)$$

と各周波数での単純混合に近似表現できる。ここで、 $H_{qk}(f)$ は信号源kからセンサqまでの周波数応答であり、 $S_k(f, \tau)$ は原信号 $s_k(f)$ に短時間離散フーリエ変換を施したものであり、 $f$ は周波数である。なお、短時間離散フーリエ変換がL個のサンプルをまとめたフレーム毎に行われるものとし、 $f_s$ をセンサのサンプリング周波数とすると、 $f$ は $f=0, f_s/L, \dots, f_s(L-1)/L$ のような離散値で表現される。また、 $\tau$ は各フレームの中心時刻（離散時間）を表現するものであり、フレームが適当な間隔（例えば、 $L/4$ ）でシフトされる場合、 $\tau$ もその間隔に従ってサンプリングされる。

【0022】

また、ベクトルを用いて式（3）を表記すると、

【数5】

$$X(f, \tau) = \sum_{k=1}^N H_k(f) \cdot S_k(f, \tau) \quad \cdots (4)$$

となる。ここで、 $X(f, \tau) = [X_1(f, \tau), \dots, X_M(f, \tau)]^T$ は混合（観測）信号ベクトル、 $H_k(f) = [H_{1k}(f), \dots, H_{Mk}(f)]^T$ は信号源kから各センサへの周波数応答をまとめたベクトルである。なお、 $[*]^T$ は $[*]$ の転置ベクトルを示す。

【0023】

〔処理の全体〕

次に、本形態の信号抽出装置1が行う信号抽出処理の全体について説明する。

図5は、本形態における信号抽出処理の全体を説明するためのフローチャートである。

以下、この図と図1とを用いて本形態における信号抽出処理の全体を説明する。

まず、前処理において、M個のセンサで観測された時間領域の混合信号 $x_q(l)$  ( $q \in \{1, \dots, M\}$ )をメモリ100の記憶領域101に格納しておく。信号抽出処理が開始されると、まず、時間周波数領域変換部120において、メモリ100の記憶領域101から時間領域の混合信号 $x_q(l)$ を読み出す。そして、時間周波数領域変換部120は、これらを短時間フーリエ変換などにより時間周波数領域の混合信号 $X_q(f, \tau)$ に変換し、メモリ100の記憶領域102に格納する（ステップS1）。

【0024】

次に、信号分離部130において、メモリ100の記憶領域102から時間周波数領域の混合信号 $X_q(f, \tau)$ を読み出す。この例の信号分離部130は、読み出した混合信号 $X_q(f, \tau)$ を要素とする混合信号ベクトル $X(f, \tau) = [X_1(f, \tau), \dots, X_M(f, \tau)]^T$ に、独立成分分析（ICA）を適用し、周波数 $f$ 毎にM行M列の分離行列 $W(f) = [W_1(f), \dots, W_M(f)]^H$ （ $*$ <sup>H</sup>は行列 $*$ の複素共役転置行列）と分離信号ベクトル

$$Y(f, \tau) = W(f) \cdot X(f, \tau) \quad \cdots (5)$$

とを算出する（ステップS2）。なお、算出された分離行列 $W(f)$ はメモリ100の記憶領

域103に格納され、分離信号ベクトル $Y(f, \tau) = [Y_1(f, \tau), \dots, Y_M(f, \tau)]^T$ を構成する分離信号 $Y_p(f, \tau)$  ( $p \in \{1, \dots, M\}$ )は記憶領域107に格納される。ステップS2の詳細については後述する。

#### 【0025】

次に、目的信号選択部140において、メモリ100の記憶領域103から分離行列 $W(f)$ を読み出し、その一般化逆行列の列である基底ベクトルを正規化し、その正規化された基底ベクトルをクラスタリングする。そして、目的信号選択部140は、そのクラスタの分散を指標として、メモリ100の記憶領域107の分離信号から目的信号 $Y_1(f)(f, \tau)$ とそれに対応する基底ベクトル $A_1(f)(f)$ とを周波数 $f$ 毎に選択し、メモリ100の記憶領域111に格納する(ステップS3)。なお、本形態では、センサに近く、そのためセンサで観測されるパワーがその他の源信号と比べて支配的であり、さらに情報として有用である(非定常性を有する)信号を目的信号として抽出する。このステップS3の詳細については後述する。

#### 【0026】

次に、時間周波数マスキング部150において、メモリ100の記憶領域102から時間周波数領域の混合信号 $X_q(f, \tau)$ を読み出し、記憶領域104から目的信号 $Y_1(f)(f, \tau)$ に対応する基底ベクトル $A_1(f)(f)$ を読み出し、これらを用いて時間周波数マスク $M(f, \tau)$ を生成して記憶領域112に格納する(ステップS4)。なお、ステップS4の処理(時間周波数マスキング部150の処理)の詳細については後述する。

#### 【0027】

次に、時間周波数マスキング部150において、メモリ100の記憶領域107から目的信号選択部140で選択された目的信号 $Y_1(f)(f, \tau)$ を読み出し、記憶領域112から時間周波数マスク $M(f, \tau)$ を読み出す。そして、時間周波数マスキング部150は、この時間周波数マスク $M(f, \tau)$ を目的信号 $Y_1(f)(f, \tau)$ に適用し、目的信号 $Y_1(f)(f, \tau)$ に残存する妨害信号成分をより一層抑圧したマスク処理目的信号 $Y_1(f)'(f, \tau)$ を生成し、メモリ100の記憶領域113に格納する(ステップS5)。なお、このステップS5の処理の詳細(時間周波数マスキング部150の処理)については後述する。

#### 【0028】

最後に、時間領域変換部160において、メモリ100の記憶領域113から時間周波数領域のマスク処理目的信号 $Y_1(f)'(f, \tau)$ を読み出し、それに短時間逆フーリエ変換(STIFT: Short-Time Inverse Fourier Transform)などを施して時間領域のマスク処理目的信号 $y(t)$ を生成し、メモリ100の記憶領域114に格納する(ステップS6)。

#### 【0029】

[ステップS2の処理(信号分離部130の処理)の詳細]

前述のように、この例の信号分離部130は、独立成分分析(ICA)によって、混合信号ベクトル $X(f, \tau) = [X_1(f, \tau), \dots, X_M(f, \tau)]^T$ から $M$ 行 $M$ 列の分離行列 $W(f) = [W_1(f), \dots, W_M(f)]^H$ と分離信号ベクトル $Y(f, \tau) = [Y_1(f, \tau), \dots, Y_M(f, \tau)]^T$ とを算出する(ステップS2)。独立成分分析(ICA)は、分離信号ベクトル $Y(f, \tau) = [Y_1(f, \tau), \dots, Y_M(f, \tau)]^T$ の各要素が互いに独立になるように分離行列 $W(f)$ を算出する手法であり、そのアルゴリズムとしては、「A. Hyvärinen and J. Karhunen and E. Oja, "Independent Component Analysis," John Wiley & Sons, 2001, ISBN 0-471-40540-X」に記載されているものなど様々なものが紹介されている。なお、独立成分分析(ICA)は、パワーが小さく、ガウス性が高く、定常な妨害信号よりも、パワーが大きく、非ガウス性が高く、非定常性が高い本形態の目的信号をより有利に分離抽出できる手法である。

#### 【0030】

[ステップS3の処理(目的信号選択部140の処理)の詳細]

独立成分分析(ICA)では、信号の独立性に着目してそれらの分離を行うため、得られる分離信号 $Y_p(f, \tau)$ には順序の任意性がある。その順序が入れ替わっても独立性が保たれるからである。従って、目的信号に対応する分離信号を周波数毎に選択しなければならない。目的信号選択部140では、以下に示す処理により、この選択を行う。

図6は、本形態における目的信号選択部140の処理の詳細を説明するためのフローチャートである。以下では、この図及び図3を用い、目的信号選択部140の処理の詳細を説明する。

#### 【0031】

まず、逆行列算出部141において、周波数毎にメモリ100の記憶領域103からM行M列の分離行列 $W(f)$ を読み出し、それらの逆行列

$$W(f)^{-1} = [A_1(f), \dots, A_M(f)] \quad (\text{ただし、その列 } A_p(f) = [A_{1p}(f), \dots, A_{Mp}(f)]^T) \quad \dots (6)$$

を算出する。ここで、式(5)の両辺に式(6)を掛け合わせることににより、時間周波数領域の混合信号 $X(f, \tau)$ の分解

#### 【数6】

$$X(f, \tau) = \sum_{p=1}^M A_p(f) Y_p(f, \tau) \quad \dots (7)$$

を得ることができる。この $A_p(f)$ は基底ベクトルと呼ばれ、それぞれが各周波数の分離信号 $Y_p(f, \tau)$ 毎に対応する。式(6)のように算出された基底ベクトル $A_p(f)$ はメモリ100の記憶領域104に格納される(ステップS11)。

#### 【0032】

次に、基底ベクトルクラスタリング部142において、全ての基底ベクトル $A_p(f)$  ( $p=1, \dots, M, f=0, f_s/L, \dots, f_s(L-1)/L$ )を正規化する。この正規化は、複数の信号源から発せられた信号の畳み込み混合を所定のモデル(例えば、近距離場モデル)に近似した場合に、正規化された基底ベクトル $A_p(f)$ が信号源の位置のみに依存したクラスタを形成するように行われる。この例では二段階の正規化によってこれを実現する。

#### 【0033】

第一の正規化は、基底ベクトルクラスタリング部142の周波数正規化部142a(図3(b))において行われる。すなわち、周波数正規化部142aにおいて、メモリ100の記憶領域104の基底ベクトル $A_p(f)$  ( $p=1, \dots, M, f=0, f_s/L, \dots, f_s(L-1)/L$ )を読み出し、これらを周波数に依存しない周波数正規化ベクトル $A_p'(f)$ に正規化してメモリ100の記憶領域105に格納する(ステップS12)。なお、この正規化は基底ベクトル $A_p(f)$ の要素 $A_{qp}(f)$ 毎に行われる(詳細は後述)。第二の正規化は、基底ベクトルクラスタリング部142のノルム正規化部142b(図3(b))において行われる。すなわち、ノルム正規化部142bにおいて、メモリ100の記憶領域105から周波数正規化ベクトル $A_p'(f)$ を読み出し、これらをノルムが規定値(この例では1)をとる正規化基底ベクトル $A_p''(f)$ に正規化してメモリ100の記憶領域106に格納する(ステップS13)。なお、この正規化は周波数正規化ベクトル $A_p'(f)$ 毎に行われる(詳細は後述)。

#### 【0034】

基底ベクトルの正規化が終了すると、次にクラスタリング部142c(図3(b))において、正規化基底ベクトル $A_p''(f)$ によって形成されるM個のクラスタ $C_i$  ( $i \in \{1, \dots, M\}$ )を同定する。この例の場合、クラスタリング部142cにおいて、メモリ100の記憶領域106から正規化基底ベクトル $A_p''(f)$ を読み出し、これらをクラスタリングしてM個のクラスタ $C_i$  ( $i=1, \dots, M$ )を生成し、各クラスタ $C_i$ を特定する情報(例えば、各クラスタに属する正規化基底ベクトル $A_p''(f)$ を示す情報)と各クラスタ $C_i$ のセントロイド(中心ベクトル) $\eta_i$ とをそれぞれメモリ100の記憶領域109, 110に格納する(ステップS14)。なお、このクラスタリングは、例えば、各クラスタ $C_i$ の要素(正規化基底ベクトル $A_p''(f)$ )と各クラスタ $C_i$ のセントロイド $\eta_i$ との間の二乗和 $U_i$ の総和 $U$

#### 【数7】

$$U = \sum_{i=1}^M U_i$$

$$U_i = \sum_{A_p''(f) \in C_i} \|A_p''(f) - \eta_i\|^2 \quad \dots (8)$$

を最小値化することを基準に行われる。この最小値化は、例えば、"R. O. Duda, P. E.

Hart, and D. G. Stork, Pattern Classification, Wiley Interscience, 2nd edition, 2000.”などで解説されているk-meansクラスタリングを用いることによって効果的に行うことができる。なお、クラスタ $C_i$ のセントロイド $\eta_i$ は、

【数8】

$$\eta_i = \frac{\sum_{A_v''(f) \in C_i} A_v''(f) / |C_i|}{\left\| \sum_{A_v''(f) \in C_i} A_v''(f) / |C_i| \right\|} \quad \dots (9)$$

によって計算される。ここで $|C_i|$ はクラスタ $C_i$ における要素（正規化基底ベクトル $A_v''(f)$ ）の数であり、 $\|\cdot\|$ はベクトル $\cdot$ のノルムである。

【0035】

M個のクラスタ $C_i$ が得られると、次に分散判定部142d（図3（b））において、目的信号に対応するクラスタを選択し、その選択情報 $l(f)$ をメモリ100の記憶領域111に格納する（ステップS15）。なお本形態では、クラスタの分散 $U_i/|C_i|$ を指標に目的信号が選択される。すなわち、本形態における基底ベクトルの正規化は、複数の信号源から発せられた信号の畳み込み混合を所定のモデルに近似した場合に、正規化された基底ベクトルが信号源の位置のみに依存したクラスタを形成するように行われる。しかし、実環境では、このようなモデルに反映されない様々な因子が存在し、このような実環境とモデルとの乖離は、信号源からセンサまでの距離が大きいほど大きくなる。例えば、近距離場モデルでは、信号が壁などに反射することによる反射波成分が考慮されておらず、直接波成分に対する反射波成分の割合は、信号源がセンサから離れるに従って大きくなる。そのため、信号がセンサから離れるに従ってモデルが成り立ちにくくなる。その結果、センサに近い信号ほど実環境に近い条件で正規化ができ、実環境とモデルとの乖離に起因するクラスタの分散を小さくできる。本形態ではセンサに近い信号を目的信号とするため、分散の小さいクラスタを目的信号に対応するクラスタとして選択すればよいことになる。なお、この選択手順（ステップS15）の詳細については後述する。

【0036】

各周波数 $f$ に対する選択情報 $l(f)$ が算出されると、次に、各周波数 $f$ における目的信号 $Y_l(f)(t)$ とそれに対応する基底ベクトル $A_l(f)(f)$ とを選択する。すなわち、まず、選択部143において、メモリ100の記憶領域111から選択情報 $l(f)$ を読み出す。そして、選択部143は、この選択情報 $l(f)$ に対応する分離信号を目的信号 $Y_l(f)(t)$ として記憶領域107から読み出し、それに対応する基底ベクトル $A_l(f)(f)$ を記憶領域104から読み出し、これらを記憶領域111に格納する（ステップS16）。

次に、ステップS12、13（図6）の正規化の詳細を示す。

【0037】

【ステップS12（第一の正規化）の詳細】

図7（a）は、ステップS12の正規化処理の詳細を説明するためのフローチャートである。

まず、制御部170（図2）においてパラメータ $p$ に1を代入し、これを一時メモリ180に格納する（ステップS21）。また、制御部170においてパラメータ $q$ に1を代入し、これを一時メモリ180に格納する（ステップS22）。次に、周波数正規化部142a（図3（b））において、メモリ100の記憶領域108から前述のパラメータ $d$ 、 $c$ 、 $Q$ を読み出し、記憶領域104から基底ベクトル $A_p(f)$ の要素 $A_{qp}(f)$ を読み出し、一時メモリ180からパラメータ $p$ 、 $q$ を読み出す。そして、この周波数正規化部142aは、基底ベクトル $A_p(f)$ の要素 $A_{qp}(f)$ に対し、

【数9】

$$A_{qp}'(f) = |A_{qp}(f)| \exp \left[ j \frac{\arg[A_{qp}(f)/A_{Qp}(f)]}{4fc^{-1}d} \right] \quad \dots (10)$$

の演算を行い、当該演算結果 $A_{qp}'(l)$ を周波数正規化ベクトル $A_p'(l)$ の各要素 $A_{qp}'(l)$ としてメモリ100の記憶領域105に格納する（ステップS23）。なお、argは偏角を意味し、expはネピア数を意味し、jは虚数単位を意味する。

【0038】

次に、制御部170において、一時メモリ180に格納されたパラメータqが $q=M$ を満たすか否かを判断する（ステップS24）。ここで $q=M$ でなければ、制御部170は、 $q+1$ の演算結果を新たなパラメータqの値とし、これを一時メモリ180に格納し（ステップS25）、処理をステップS23へ戻す。一方、 $q=M$ であれば、制御部170は、さらに $p=N$ を満たすか否かを判断する（ステップS26）。

ここで $p=N$ でなければ、制御部170において、 $p+1$ の演算結果を新たなパラメータpの値とし、これを一時メモリ180に格納し（ステップS27）、処理をステップS22へ戻す。一方、 $p=N$ であれば、制御部170は、ステップS12の処理を終了させる（ステップS12（第一の正規化）の詳細の説明終わり）。

【0039】

〔ステップS13（第二の正規化）の詳細〕

図7（b）は、ステップS13の正規化処理の詳細を説明するためのフローチャートである。

まず、制御部170においてパラメータpに1を代入し、これを一時メモリ180に格納する（ステップS31）。次に、ノルム正規化部142bにおいて、メモリ100の記憶領域105から周波数正規化ベクトル $A_p'(l)$ の各要素 $A_{qp}'(l)$ を読み出し、

【数10】

$$\|A_p'(f)\| = \sqrt{\sum_{q=1}^M (A_{pq}'(f))^2} \quad \dots (11)$$

の演算を行って、周波数正規化ベクトル $A_p'(l)$ のノルム $\|A_p'(l)\|$ を求め、周波数正規化ベクトル $A_p'(l)$ 及びそのノルム $\|A_p'(l)\|$ を一時メモリ180に格納する（ステップS32）。

【0040】

次に、ノルム正規化部142bにおいて、一時メモリ180から周波数正規化ベクトル $A_p'(l)$ 及びそのノルム $\|A_p'(l)\|$ を読み出し、

$$A_{p''}'(l) = A_p'(l) / \|A_p'(l)\| \quad \dots (12)$$

の演算を行って正規化基底ベクトル $A_{p''}'(l)$ を求め、これをメモリ100の記憶領域106に格納する（ステップS33）。その後、制御部170において、一時メモリ180に格納されたパラメータpが $p=N$ を満たすか否かを判断する（ステップS34）。ここで $p=N$ でなければ、制御部170は、 $p+1$ の演算結果を新たなパラメータpの値とし、これを一時メモリ180に格納し（ステップS35）、ステップS32の処理に戻す。一方、 $p=N$ であれば、制御部170は、ステップS13の処理を終了させる。なお、正規化基底ベクトル $A_{p''}'(l)$ がクラスタを形成する理由については後述する（ステップS13（第二の正規化）の詳細の説明終わり）。

このように生成された正規化基底ベクトル $A_{p''}'(l)$ は、周波数に依存せず、信号源の位置のみに依存するベクトルとなる。その結果、この正規化基底ベクトル $A_{p''}'(l)$ はクラスタを形成することになる。以下にこの理由を説明する。

【0041】

〔正規化基底ベクトル $A_{p''}'(l)$ がクラスタを形成する理由〕

まず、前述した式（2）に示される畳み込み混合モデルを直接波（近距離場）混合モデル

【数 1 1】

$$h_{qk}(f) = \frac{\phi(f)}{d_{qk}} \exp[-j2\pi f c^{-1}(d_{qk} - d_{Qk})] \quad \dots (13)$$

を用いて近似する。ここで、 $\phi(f)$ は信号の減衰率、 $d_{qk}$ は信号源 $q$ とセンサ $k$ との間の距離であり、また、減衰 $1/d_{qk}$ は、距離 $d_{qk}$ によって決まり、遅延 $(d_{qk} - d_{Qk})/c$ は、基準値 $Q$ に対応する基準センサ $Q$ の位置で正規化された距離によって決まる。

【0 0 4 2】

また、独立信号分離(ICA)における順序の任意性とスケーリングの任意性を考慮すると、基底ベクトル $A_p(f)$ と式(4)における信号源 $k$ から各センサへの周波数応答をまとめたベクトル $H_k(f)$ との間には、以下の関係が成り立つ。

$$A_p(f) = \varepsilon_p \cdot H_k(f), \quad A_{qp}(f) = \varepsilon_p \cdot H_{qk}(f) \quad \dots (14)$$

ここで、 $\varepsilon_p$ はスケーリングの任意性を表現する複素スカラー値であり、添え字 $k$ と $q$ が異なる可能性が順序の任意性を表現している。

【0 0 4 3】

式(13)と(14)を式(10)と(12)に代入すると、

【数 1 2】

$$A_{qp}''(f) = \frac{1}{d_{qk}D} \exp\left[-j\frac{\pi(d_{qk} - d_{Qk})}{d}\right], \quad D = \sqrt{\sum_{i=1}^M \frac{1}{d_{ik}^2}} \quad \dots (15)$$

となる。この式から分かるように、正規化基底ベクトル $A_p''(f)$ の各要素 $A_{qp}''(f)$ は、周波数 $f$ からは独立であり、信号源 $k$ とセンサ $j$ の位置のみに依存する。従って、正規化基底ベクトル $A_p''(f)$ をクラスタリングすると、同じ信号源にごとに対応するクラスが形成される。

【0 0 4 4】

同様なことは、信号の減衰を考慮していない近距離場混合モデルでモデル化した場合にも言える。すなわち、前述した式(2)に示される畳み込み混合モデルを、減衰を無視した近距離場混合モデルで近似すると

【数 1 3】

$$h_{qk}(f) = \exp[-j2\pi f c^{-1}(d_{qk} - d_{Qk})] \quad \dots (16)$$

となる。そして、式(16)と(14)を式(10)と(12)に代入すると、

【数 1 4】

$$A_{qp}''(f) = \frac{1}{\sqrt{M}} \exp\left[-j\frac{\pi(d_{qk} - d_{Qk})}{d}\right] \quad \dots (17)$$

となる。この場合も、正規化基底ベクトル $A_p''(f)$ の各要素 $A_{qp}''(f)$ は、周波数 $f$ からは独立であり、信号源 $k$ とセンサ $j$ の位置のみに依存している。

【0 0 4 5】

さらに、近距離場混合モデルだけではなく、遠距離場混合モデルでも同様なことがいえる。すなわち、前述した式(2)に示される畳み込み混合モデルを遠距離場混合モデルで近似すると、

【数 1 5】

$$h_{qk}(f) = \exp[-j2\pi f c^{-1} \|SE_q - SE_Q\| \cos\theta_k^{qQ}] \quad \dots (18)$$

となる。なお、 $SE_q$ 及び $SE_Q$ はセンサ $q$ 、 $Q$ の位置を示すベクトルである。また $\theta_k^{qQ}$ は、

センサ  $q$ ,  $Q$  を結ぶ直線と、センサ  $q$ ,  $Q$  の中心点と信号源  $k$  とを結ぶ直線と、で成す角度を示す。そして、式 (18) と (14) を式 (10) と (12) に代入すると、

【数 16】

$$A_{qp}''(f) = \frac{1}{\sqrt{M}} \exp \left[ -j \frac{\pi \|SE_q - SE_Q\| \cos \theta_k^{qQ}}{d} \right] \quad \dots (19)$$

となる。この場合も、正規化基底ベクトル  $A_p''(f)$  の各要素  $A_{qp}''(f)$  は、周波数  $f$  からは独立であり、信号源  $k$  とセンサ  $j$  の位置のみに依存している。

また、式 (15) から、式 (10) におけるパラメータ  $d$  の値は、 $d > d_{\max}/2$  であることが望ましく ( $d_{\max}$  は要素  $A_{Qp}(f)$  に対応する基準センサ  $Q$  と他のセンサとの最大距離  $d_{\max}$  を意味する。)、より好ましくは  $d \geq d_{\max}$  であることが望ましく、さらにより好ましくは  $d = d_{\max}$  であることが望ましいことが分かる。以下、この理由について説明する。

【0046】

図 9 及び図 10 は、パラメータ  $d$  毎の正規化基底ベクトルの要素  $A_{qp}''(f)$  とその偏角  $\arg[A_{qp}''(f)]$  との関係を示すための複素平面図である。なお、これらにおける横軸は実軸を縦軸は虚軸を示している。

図 9 (a) は  $d_{\max}/2 \geq d$  の場合における複素平面図である。ここで上述の  $d_{\max}$  の定義より、任意の  $q$  及び  $k$  に対し  $d_{qk} - d_{Qk}$  の絶対値は  $d_{\max}$  以下となる。よって  $d_{\max}/2 \geq d$  の場合、 $(\pi/2) \cdot (d_{qk} - d_{Qk})/d \leq -\pi$ ,  $(\pi/2) \cdot (d_{qk} - d_{Qk})/d \geq \pi$  と成り得る。その結果、式 (15) で表される  $A_{qp}''(f)$  の偏角  $\arg[A_{qp}''(f)]$  は、 $360^\circ$  を超える  $\alpha_1 \leq \arg[A_{qp}''(f)] \leq \alpha_2$  ( $\alpha_1 \leq -180^\circ$ ,  $\alpha_2 \geq 180^\circ$ ) の範囲に分布する可能性がある。そのため、異なる正規化基底ベクトルの要素  $A_{qp}''(f)$  の偏角が一致する可能性があり、前述したクラスタリングにおいて、異なる正規化基底ベクトル  $A_p''(f)$  を同じクラスにクラスタリングしてしまう可能性がある。よって、 $d > d_{\max}/2$  であることが望ましい。しかし、この偏角の重複範囲に対応する正規化基底ベクトル  $A_p''(f)$  のサンプルが存在しないならば  $d_{\max}/2 \geq d$  としても問題はない。

【0047】

図 9 (b) は  $d_{\max}/2 < d < d_{\max}$  の場合における複素平面図である。この場合、 $-\pi < (\pi/2) \cdot (d_{qk} - d_{Qk})/d < -\pi/2$ ,  $\pi/2 < (\pi/2) \cdot (d_{qk} - d_{Qk})/d < \pi$  と成り得る。その結果、式 (15) で表される  $A_{qp}''(f)$  の偏角  $\arg[A_{qp}''(f)]$  は、 $\beta_1 \leq \arg[A_{qp}''(f)] \leq \beta_2$  ( $-180^\circ < \beta_1 < -90^\circ$ ,  $90^\circ < \beta_2 < 180^\circ$ ) の範囲に分布する可能性がある。そのため、 $-180^\circ < \arg[A_{qp}''(f)] < -90^\circ$  及び  $90^\circ < \arg[A_{qp}''(f)] < 180^\circ$  の範囲において、異なる周波数正規化ベクトル要素間の偏角の差の増加に伴い、これらの周波数正規化ベクトル要素間の距離が単調増加しないこともありうる。これは、前述したクラスタリングの精度を低下させる可能性がある。よって  $d \geq d_{\max}$  であることがより望ましい。

【0048】

図 10 (a) は  $d = d_{\max}$  の場合における複素平面図であり、図 10 (b) は  $d > d_{\max}$  の場合における複素平面図である。ここで  $d > d_{\max}$  の場合、 $-\pi/2 < (\pi/2) \cdot (d_{qk} - d_{Qk})/d < 0$ ,  $0 < (\pi/2) \cdot (d_{qk} - d_{Qk})/d < \pi/2$  と成り得る。その結果、式 (15) で表される  $A_{qp}''(f)$  の偏角  $\arg[A_{qp}''(f)]$  は、図 10 (b) に示すように、 $\gamma_1 \leq \arg[A_{qp}''(f)] \leq \gamma_2$  ( $-90^\circ < \gamma_1 < 0^\circ$ ,  $0^\circ < \gamma_2 < 90^\circ$ ) の範囲に分布する。そして、 $d$  が大きくなればなるほどその分布範囲は狭くなっていき、狭い範囲にクラスが密集していく。これは、前述したクラスタリングの精度を低下させる。

【0049】

これに対し、 $d = d_{\max}$  である場合、 $-\pi/2 \leq (\pi/2) \cdot (d_{qk} - d_{Qk})/d < 0$ ,  $0 < (\pi/2) \cdot (d_{qk} - d_{Qk})/d \leq \pi/2$  と成り得る。その結果、式 (15) で表される  $A_{qp}''(f)$  の偏角  $\arg[A_{qp}''(f)]$  は図 10 (a) に示すように  $-\pi/2 \leq \arg[A_{qp}''(f)] \leq \pi/2$  の範囲に分布する。この場



合、周波数正規化ベクトルの要素間における偏角の差の増加に対して、それらの距離も単調増加するという関係を維持しつつ、できるだけ広い範囲にクラスタを分散させることができる。その結果、一般的にクラスティングの精度を向上させることができる（正規化基底ベクトル $A_p''(l)$ がクラスタを形成する理由の説明終わり）。

#### 【0050】

【目的信号の選択手順（ステップS15）の詳細】

次に、前述した目的信号の選択手順（ステップS15）の詳細を例示する。

クラスタの選択手順1：

第1の例は、最も分散が小さいクラスタを目的信号に対応するクラスタとして選択するものである。図8は、この第1の例を説明するためのフローチャートである。

この場合、まず分散判定部142d（図3（b））において、メモリ100の記憶領域109からクラスタ $C_i$ （ $i \in \{1, \dots, M\}$ ）を特定する情報を読み出し、さらに記憶領域106、110から、それぞれ正規化基底ベクトル $A_p''(l) \in C_i$ 及びセントロイド $\eta_i$ を読み出す。そして、分散判定部142dは、各 $i$ について、式（8）のように $U_i$ を求め、 $C_i$ に属する要素（正規化基底ベクトル $A_p''(l)$ ）の数をカウントすることにより $|C_i|$ を求め、クラスタ $C_i$ の分散 $U_i/|C_i|$ を算出して一時メモリ180に格納する。そして、分散判定部142dは、この一時メモリ180に格納された分散 $U_i/|C_i|$ から最小なものを選択し、それに対応するクラスタを示す情報をクラスタ選択情報

$$i = \operatorname{argmin}_i U_i/|C_i| \quad \dots (20)$$

として一時メモリ180に格納する（ステップS41）。なお、式（20）における $\operatorname{argmin}_i$ は、 $i$ を最小にする $i$ を意味する。

#### 【0051】

次に、制御部170（図2）において、パラメータ $f$ に0を代入し、これを一次メモリ180に格納する（ステップS42）。

次に、分散判定部142dは、一時メモリ180からクラスタ選択情報 $i$ を読み出し、メモリ100の記憶領域110からクラスタ選択情報 $i$ に対応するセントロイド $\eta_i$ を読み出す。また、分散判定部142dは、メモリ100の記憶領域106から正規化基底ベクトル $A_p''(l)$ （ $p \in \{1, \dots, M\}$ ）を読み出す。そして、分散判定部142dは、周波数 $l$ 毎に選択情報

$$l(l) = \operatorname{argmin}_p \|A_p''(l) - \eta_i\|^2 \quad \dots (21)$$

を算出し、記憶領域111に格納する（ステップS43）。

#### 【0052】

次に、制御部170において、一時メモリ180からパラメータ $f$ を読み出し、 $f = (l-1) \cdot f_s/L$ であるか否かを判断する（ステップS44）。ここで $f = (l-1) \cdot f_s/L$ でなければ、制御部170は、このパラメータ $f$ に $f_s/L$ を加算した値を新たな $f$ として一時メモリ180に格納し（ステップS45）、処理をステップS43に戻す。一方、 $f = (l-1) \cdot f_s/L$ であれば、制御部170は、ステップS15の処理を終了させる。

クラスタの選択手順2：

第2の例は、予め定められたしきい値よりも分散が小さいクラスタを目的信号に対応するクラスタとして選択するものである。なお、このしきい値は、例えば実験結果又は経験に基づき定められる値であり、事前にメモリ100に格納しておくものである。

#### 【0053】

この場合、ステップS41（図8）の処理の代わりに、分散判定部142dにおいて、まず各クラスタの分散 $U_i/|C_i|$ を公知の並び替えアルゴリズムによって昇順或いは降順に並び替える。そして、分散判定部142dは、メモリ100に格納されたしきい値を読み取り、分散 $U_i/|C_i|$ がしきい値よりも小さくなるクラスタを選択し、選択したクラスタに対応する添字 $i$ をクラスタ選択情報 $i$ として一時メモリ180に格納する。それ以外は、クラスタの選択手順1と同様である。

クラスタの選択手順3：

第3の例は、最も分散の小さいクラスタだけではなく、周波数毎に、最も分散が小さい

クラスタから所定番目に分散が小さいクラスタまで（例えば、最も分散の小さいクラスタから3番目に分散が小さいクラスタまで）を目的信号に対応するクラスタとして選択するものである。

【0054】

この場合、ステップS46（図8）の処理の代わりに、分散判定部142dにおいて、まず各クラスタの分散 $U_i/|C_i|$ を公知の並び替えアルゴリズムによって昇順或いは降順に並び替える。そして、分散判定部142dは、最も分散 $U_i/|C_i|$ が小さいクラスタから所定番目に分散 $U_i/|C_i|$ が小さいクラスタまでを選択し、選択したクラスタに対応する、選択したクラスタに対応する添字 $i$ をクラスタ選択情報 $i$ として一時メモリ180に格納する。それ以外は、クラスタの選択手順1と同様である。

なお、クラスタの選択手順1の代わりに分散の小ささが2番目以降のいずれかのクラスタを選択することとしてもよく、上述したクラスタの選択手順の一部を組み合わせた手順を適用してもよい（ステップS15の説明終わり／ステップS3の処理（目的信号選択部140の処理）の詳細の説明終わり）。

【0055】

【時間周波数マスクング部150の処理（ステップS4，S5）の詳細】

次に、時間周波数マスクング部150の処理について説明する。前述のように、時間周波数マスクング部150は、目的信号選択部140で選択された目的信号 $Y_1(f)(f, \tau)$ に残存する妨害信号成分をさらに抑圧するものである。以下ではまず、なぜ目的信号 $Y_1(f)(f, \tau)$ に妨害信号が残存するのかについて説明する。

前述した式（5）を目的信号のみに関して書くと、

$$Y_1(f)(f, \tau) = W_1(f)^H \cdot H(f) \cdot X(f, \tau) \quad \cdots (22)$$

となる。また、これに式（4）を代入して周波数 $f$ を省略して書くと、

【0056】

【数17】

$$Y_1(\tau) = W_1^H \cdot H_1 \cdot S_1(\tau) + \sum_{k=2}^N W_1^H \cdot H_k \cdot S_k(\tau) \quad \cdots (23)$$

と表記できる。もし $N \leq M$ であれば、独立成分分析（ICA）により、 $W_1^H \cdot H_k = 0, \forall k \in \{2, \dots, N\}$ を満たすような $W_1$ を設定することが可能である。そのため、式（23）の第二項は0になる。しかし、より一般的に信号源数 $N$ がセンサ数 $M$ よりも多い場合には、 $W_1^H \cdot H_k \neq 0, \forall k \in \kappa$ となる $\kappa \subseteq \{2, \dots, N\}$ が存在してしまう。その場合、目的信号 $Y_1(f)$ は unnecessary 残留成分（妨害信号の残留成分）

【数18】

$$\sum_{k \in \kappa} W_1^H \cdot H_k \cdot S_k(\tau)$$

を含んでしまう（以後 $f$ を省略せず）。

【0057】

時間周波数マスクング部150の目的は、この目的信号 $Y_1(f, \tau)$ の unnecessary 残留成分を抑圧し、目的信号 $Y_1(f, \tau)$ よりも妨害信号の残留成分が少ないマスク処理目的信号 $Y_1'(f, \tau)$ を生成することである。そのために、時間周波数マスクング部150は、まずマスク生成部151（図4）において、妨害信号の残留成分を有する時間周波数スロットに対して小さな値を採り、妨害信号の残留成分が少ない時間周波数スロットに対して大きな値を採る時間周波数マスク $0 \leq M(f, \tau) \leq 1$ を生成する。そして、マスクング部152において、

$$Y_1(f)'(f, \tau) = M(f, \tau) \cdot Y_1(f)(f, \tau) \quad \cdots (24)$$

に従ってマスクング処理を行い、マスク処理目的信号 $Y_1(f)'(f, \tau)$ を出力する。以下、こ

の詳細について説明する。

#### 【0058】

【ステップS4（マスク生成部151）の処理の詳細】

図11は、図5におけるステップS4の詳細を説明するためのフローチャートである。以下、この図と図4とを用い、ステップS4（マスク生成部151）の処理の詳細を説明する。

この例のマスク生成部151は、時間周波数領域の混合信号ベクトル $X(l, \tau)$ を白色化する空間（白色化空間）において、混合信号ベクトル $X(l, \tau)$ と目的信号に対応する基底ベクトル $A_l(l)$ とのなす角度 $\theta_l(l)(l, \tau)$ を求め、その角度 $\theta_l(l)(l, \tau)$ に基づき、時間周波数マスクを生成する。なお、白色化とは、混合信号ベクトル $X(l, \tau)$ を線形に変形し、その共分散行列が単位行列と等しくなるようにすることである。

#### 【0059】

そのためにまず、白色化行列生成部151aにおいて、時間周波数領域の混合信号 $X_q(l, \tau)$ を用い、混合信号ベクトル $X(l, \tau)$ を白色化空間に移す白色化行列 $V(l)$ を生成する（ステップS51）。この例の場合、白色化行列生成部151aは、メモリ100の記憶領域102から混合信号 $X_q(l, \tau)$ を読み出し、 $R(l) = \langle X(l, \tau) \cdot X(l, \tau)^H \rangle_\tau$ とした場合における $V(l) = R(l)^{-1/2}$ を白色化行列 $V(l)$ として算出し、これを記憶領域112に格納する。なお、 $\langle * \rangle_\tau$ はベクトル $*$ を時間平均したベクトルを意味し、 $*$ <sup>H</sup>は $*$ の複素共役転置を意味し、 $R^{-1/2}$ は、 $V(l) \cdot R(l) \cdot V(l)^H = I$ （ $I$ は単位行列）を満たす行列を意味する。また、白色化行列 $V(l)$ を算出する代表的な方法として、 $R(l)$ を固有値分解 $R(l) = E(l) \cdot D(l) \cdot E(l)^H$ （ $E(l)$ はユニタリ行列、 $D(l)$ は対角行列）し、 $V(l) = D(l)^{-1/2} \cdot E(l)^H$ の演算を行う方法を例示できる。なお、 $D(l)^{-1/2}$ は、対角行列 $D(l)$ の各要素の $(-1/2)$ 乗した対角行列と等価となるため、対角行列 $D(l)$ の各要素の $(-1/2)$ 乗することによって算出できる。

#### 【0060】

次に、白色化部151bにおいて、この白色化行列 $V(l)$ を用い、混合信号ベクトル $X(l, \tau)$ を白色化空間に写像した白色化混合信号ベクトル $Z(l, \tau)$ と、基底ベクトル $A_l(l)$ を白色化空間に写像した白色化基底ベクトル $B_l(l)$ とを算出する（ステップS52）。この例の場合、まず白色化部151bにおいて、メモリ100の記憶領域102から混合信号 $X_q(l, \tau)$ を、記憶領域111から目的信号 $Y_l(l)(l, \tau)$ に対応する基底ベクトル $A_l(l)(l)$ を、記憶領域112から白色化行列 $V(l)$ をそれぞれ読み出す。そして、白色化部151bは、

$Z(l, \tau) = V(l) \cdot X(l, \tau)$ の演算により白色化混合信号ベクトル $Z(l, \tau)$ を算出し、 $B_l(l)(l) = V(l) \cdot A_l(l)(l)$ の演算により白色化基底ベクトル $B_l(l)(l)$ を算出し、これらをメモリ100の記憶領域112に格納する。

#### 【0061】

次に、角度計算部151cにおいて、白色化混合信号ベクトル $Z(l, \tau)$ と白色化基底ベクトル $B_l(l)(l)$ とかなす角度 $\theta_l(l)(l, \tau)$ を時間周波数毎に算出する（ステップS53）。この例の場合、まず角度計算部151cにおいて、メモリ100の記憶領域112から白色化混合信号ベクトル $Z(l, \tau)$ と白色化基底ベクトル $B_l(l)(l)$ とを読み出す。そして、角度計算部151cは、

$$\theta_l(l)(l, \tau) = \cos^{-1}(|B_l(l)(l)^H \cdot Z(l, \tau)| / \|B_l(l)(l)\| \cdot \|Z(l, \tau)\|) \quad \dots (25)$$

の演算により、各時間周波数スロットに対して角度 $\theta_l(l)(l, \tau)$ を算出し、記憶領域112に格納する。なお、式(25)における $|*|$ はベクトル $*$ の絶対値を意味し、 $\|*\|$ はベクトル $*$ のノルムを意味する。

#### 【0062】

次に、関数演算部151dにおいて、角度 $\theta_l(l)(l, \tau)$ を要素とする関数である時間周波数マスク $M(l, \tau)$ を生成する（ステップS54）。この例の場合、関数演算部151dにおいて、まず、メモリ100の記憶領域108から実数パラメータ $\theta_T, g$ を、記憶領域112から角度 $\theta_l(l)(l, \tau)$ をそれぞれよみだす。そして、関数演算部151dは、ロジスティック関数

$$M(\theta(f, \tau)) = 1 / (1 + e^g \cdot (\theta(f, \tau) - \theta_T)) \quad \dots (26)$$

を時間周波数マスク $M(f, \tau)$ として算出する。なお、実数パラメータ $\theta_T, g$ は、それぞれ時間周波数マスク $M(f, \tau)$ の転移点と傾き度を指定するパラメータである。図12(a)に、二種類の実数パラメータ $\theta_T, g$ に対し、式(26)に従って算出された時間周波数マスク $M(f, \tau)$ を例示する。この図に例示するように、実数パラメータ $\theta_T$ が小さいほど、時間周波数マスク $M(f, \tau)$ が大きな値(この例では1)をとる領域が狭くなる。これは実数パラメータ $\theta_T$ が小さいほど、マスク処理目的信号 $Y_I(f)'(f, \tau)$ に残存する妨害信号成分が少なくなるが、その反面、マスク処理目的信号 $Y_I(f)'(f, \tau)$ が不自然なものとなるといった傾向に現れる。例えば、音声信号を目的信号とする場合、実数パラメータ $\theta_T$ を小さくするほどミュージカルノイズ(musical noise)が大きくなる。また、実数パラメータ $g$ が大きくなるほど時間周波数マスク $M(f, \tau)$ の波形(大きな値(この例では1)から小さな値(この例では0)への遷移)は急峻になる。ここで、できるだけマスク処理目的信号 $Y_I(f)'(f, \tau)$ に残存する妨害信号成分を減らし、マスク処理目的信号 $Y_I(f)'(f, \tau)$ を自然な信号とするためには実数パラメータ $g$ を小さくし、時間周波数マスク $M(f, \tau)$ の波形を滑らかなものとするのが望ましい。

#### 【0063】

なお、これらの実数パラメータ $\theta_T, g$ を周波数毎に設定しておく構成としてもよい。また、さらに実数パラメータ $\alpha$ を導入し、ロジスティック関数

$$M(\theta(f, \tau)) = \alpha / (1 + e^g \cdot (\theta(f, \tau) - \theta_T)) \quad \dots (27)$$

を時間周波数マスク $M(f, \tau)$ としてもよい。さらに、角度 $\theta_I(f)'(f, \tau)$ が0に近い領域において大きな値を採り、角度 $\theta_I(f)'(f, \tau)$ が大きな領域において小さな値を採る $0 \leq M(f, \tau) \leq 1$ の関数であれば、他の関数を時間周波数マスク $M(f, \tau)$ として用いてもよい(ステップS4(マスク生成部151)の処理の詳細の説明終わり)。

#### 【0064】

【ステップS5(マスキング部152)の処理の詳細】

マスキング部152は、メモリ100の記憶領域111から目的信号 $Y_I(f)'(f, \tau)$ を、記憶領域112から時間周波数マスク $M(f, \tau)$ をそれぞれ読み出す。そして、マスキング部152は、

$$Y_I(f)'(f, \tau) = M(f, \tau) \cdot Y_I(f)'(f, \tau) \quad \dots (28)$$

の演算によってマスク処理目的信号 $Y_I(f)'(f, \tau)$ を求め、これをメモリ100の記憶領域113に格納する(ステップS5(マスキング部152)の処理の詳細の説明終わり)。

#### 【0065】

【時間周波数マスキングの効果】

次に、上述した時間周波数マスク $M(f, \tau)$ によるマスキングの効果について説明する。

信号源のスパース性が高く、信号源 $S_k(f, \tau)$ がゼロに近くなる可能性が高いなら、式(4)は、

#### 【数19】

$$X(f, \tau) \approx H_k(f) \cdot S_k(f, \tau), k \in \{1, \dots, N\} \quad \dots (29)$$

として近似される。なお、式(29)における $k$ は各信号源に対応する添字であり、時間周波数位置 $(f, \tau)$ 毎に決まる。よって、目的信号が唯一アクティブ或いはそれに近い時間周波数位置 $(f, \tau)$ では、白色化混合信号ベクトル $Z(f, \tau)$ は、

#### 【数20】

$$Z(f, \tau) \approx V(f) \cdot H_{I(f)}(f) \cdot S_{I(f)}(f, \tau) \approx V(f) \cdot A_{I(f)}(f) \cdot Y_{I(f)}(f, \tau)$$

と近似できる。ここで、 $Y_I(f)'(f, \tau)$ はスカラー値である。また、前述のように白色化基底ベクトル $B_I(f)'(f)$ は、

$$B_I(f)'(f) = V(f) \cdot A_I(f)'(f) \quad \dots (30)$$

である。以上より、目的信号が唯一アクティブ或いはそれに近い時間周波数位置 $(f, \tau)$ では、白色化混合信号ベクトル $Z(f, \tau)$ と白色化基底ベクトル $B_I(f)'(f)$ とがなす角度 $\theta_I(f)'(f)$

$l, \tau)$ は0に近くなることから分かる。そして、前述のように時間周波数マスク $M(l, \tau)$ は、角度 $\theta_{|l|}(l, \tau)$ が0に近い領域において大きな値を採るのである。よって、時間周波数マスク $M(l, \tau)$ は、目的信号が唯一アクティブ或いはそれに近い時間周波数位置 $(l, \tau)$ における目的信号 $Y_{|l|}(l, \tau)$ をマスク処理目的信号 $Y_{|l|}(l, \tau)$ として抽出するものといえる(式(28)参照)。

【0066】

一方、仮に $l(l)=1$ とおいた場合、目的信号が殆どアクティブでない時間周波数位置 $(l, \tau)$ では、白色化混合信号ベクトル $Z(l, \tau)$ は、

【数21】

$$Z(f, \tau) \approx \sum_{k=2}^N V(f) \cdot H_k(f) \cdot S_k(f, \tau) \quad \cdots (31)$$

と近似できる。ここで、もし信号源の $N$ が、センサの数 $M$ と同等かそれより少ない場合、白色化空間でのベクトル $V(l) \cdot H_1(l), \dots, V(l) \cdot H_k(l)$ は互いに直交する。ここで、式(31)の $S_k(l, \tau)$ はスカラー値である。これより、白色化混合信号ベクトル $Z(l, \tau)$ と白色化基底ベクトル $B_{|l|}(l)$ とがなす角度 $\theta_{|l|}(l, \tau)$ は大きい角度を示すことがわかる。また $N > M$ でも、白色化基底ベクトル $B_{|l|}(l)$  ( $l(l)=1$ )は、目的信号以外のベクトル $V(l) \cdot H_2(l), \dots, V(l) \cdot H_k(l)$ とは大きい角度を持つ傾向がある。以上より、目的信号が殆どアクティブでない時間周波数位置 $(l, \tau)$ では、角度 $\theta_{|l|}(l, \tau)$ が大きな値を採ることが分かる。そして、時間周波数マスク $M(l, \tau)$ は、角度 $\theta_{|l|}(l, \tau)$ が0から離れた領域において小さな値を採るのであるから、目的信号が殆どアクティブでない時間周波数位置 $(l, \tau)$ における目的信号 $Y_{|l|}(l, \tau)$ をマスク処理目的信号 $Y_{|l|}(l, \tau)$ から排除するものといえる(式(28)参照)。

【0067】

以上より、時間周波数マスク $M(l, \tau)$ による時間周波数マスキング処理は、目的信号 $Y_{|l|}(l, \tau)$ に残存する妨害信号成分をより一層抑圧するものであることがわかる。

なお、この時間周波数マスキング処理は、音声や音楽などスパース性を持つ信号に対して特に有効である。スパース性が低い場合、目的信号がアクティブである時間周波数位置 $(l, \tau)$ であっても他の妨害信号成分も多く含まれ、式(29)の近似が成り立たなくなり、角度 $\theta_{|l|}(l, \tau)$ が0から離れてしまうからである。すなわちスパースでない場合、例えば、図12(b)のように、ある時間周波数位置 $(l, \tau)$ において目的信号( $l(l)=1$ とする)に対応するベクトル $V(l) \cdot H_1(l)$ の他に、妨害信号に対応ベクトルする $V(l) \cdot H_2(l)$ 、 $V(l) \cdot H_3(l)$ が共存することになる。この例の場合、白色化混合信号ベクトル $Z(l, \tau)$ は、

【数22】

$$Z(f, \tau) \approx \sum_{k=1}^3 V(f) \cdot H_k(f) \cdot S_k(f, \tau) \quad \cdots (32)$$

となるため、白色化混合信号ベクトル $Z(l, \tau)$ と白色化基底ベクトル $B_{|l|}(l)$ とがなす角度 $\theta_{|l|}(l, \tau)$ も0から離れてしまう。これは目的信号がアクティブである時間周波数位置 $(l, \tau)$ の信号をマスク処理目的信号 $Y_{|l|}(l, \tau)$ から排除してしまう可能性があることを示している。

また、この時間周波数マスキング処理は、目的信号のパワーが妨害信号のパワーに比べて十分に大きい場合に特に有効である。すなわち、目的信号のパワーが妨害信号のパワーに比べて十分に大きい場合、たとえスパース性が低く、目的信号がアクティブである時間周波数位置 $(l, \tau)$ に他の妨害信号成分も含まれる状況であっても、式(29)の近似が比較的成り立ち、角度 $\theta_{|l|}(l, \tau)$ が0に近づくからである。例えば、目的信号のパワーが妨害信号のパワーに比べて十分に大きい場合例えば式(32)における妨害信号の寄与度は低くなり、白色化混合信号ベクトル $Z(l, \tau)$ と白色化基底ベクトル $B_{|l|}(l)$ とがなす角度 $\theta_{|l|}(l, \tau)$ も0に近づく。これは目的信号がアクティブである時間周波数位置 $(l, \tau)$ の信号をマスク処理目的信号 $Y_{|l|}(l, \tau)$ から排除してしまう可能性が低くなり、またマスク処理目的信号 $Y_{|l|}(l, \tau)$ に尚も残存する妨害信号成分を相対的に低くできることを意味している(ステップS5(マスキング部152)の処理の詳細の説明終わり)。

## 【0068】

### （第2の実施の形態）

次に、本発明における第2の実施の形態について説明する。

本形態は、第1の実施の形態の変形例であり、時間周波数マスクによる処理を行わない点のみが第1の実施の形態と相違する。以下では、第1の実施の形態との相違点を中心に説明し、第1の実施の形態と共通する事項については説明を省略する。

#### ＜構成＞

図13は、CPU10（図1）に信号抽出プログラムが読み込まれることにより構成される本形態における信号抽出装置200のブロック図の例示である。

## 【0069】

図13に例示するように、本形態の信号抽出装置200の信号抽出装置1との相違点は、メモリ100に記憶領域112、113を有しない点、及び時間周波数マスクング部150を有しない点である。

#### ＜処理＞

次に、本形態における信号抽出装置200の処理について説明する。

図14は、本形態における信号抽出装置200の処理を説明するためのフローチャートである。以下では、第1の実施の形態との相違点を中心に説明する。

## 【0070】

まず、第1の実施の形態と同様に、時間周波数領域変換部120において、メモリ100の記憶領域101から時間領域の混合信号 $X_q(l)$ を読み出す。そして、時間周波数領域変換部120は、これらを短時間フーリエ変換などにより時間周波数領域の混合信号 $X_q(l, \tau)$ に変換し、メモリ100の記憶領域102に格納する（ステップS61）。

次に、信号分離部130において、メモリ100の記憶領域102から時間周波数領域の混合信号 $X_q(l, \tau)$ を読み出す。この例の信号分離部130は、読み出した混合信号 $X_q(l, \tau)$ を要素とする混合信号ベクトル $X(l, \tau) = [X_1(l, \tau), \dots, X_M(l, \tau)]^T$ に、独立成分分析（ICA）を適用し、周波数 $f$ 毎に $M$ 行 $M$ 列の分離行列 $W(l) = [W_1(l), \dots, W_M(l)]^H$ （ $*$ <sup>H</sup>は行列 $*$ の複素共役転置行列）と分離信号ベクトル $Y(l, \tau) = W(l) \cdot X(l, \tau)$ とを算出する（ステップS62）。なお、算出された分離行列 $W(l)$ はメモリ100の記憶領域103に格納され、分離信号ベクトル $Y(l, \tau) = [Y_1(l, \tau), \dots, Y_M(l, \tau)]^T$ を構成する分離信号 $Y_p(l, \tau)$ （ $p \in \{1, \dots, M\}$ ）は記憶領域107に格納される。

## 【0071】

次に、目的信号選択部140において、メモリ100の記憶領域103から分離行列 $W(l)$ を読み出し、その一般化逆行列の列である基底ベクトルを正規化し、その正規化された基底ベクトルをクラスタリングする。そして、目的信号選択部140は、そのクラスタの分散を指標として、メモリ100の記憶領域107の分離信号から目的信号 $Y_1(l) (l, \tau)$ を周波数 $f$ 毎に選択し、メモリ100の記憶領域111に格納する（ステップS63）。

次に、時間領域変換部160において、メモリ100の記憶領域111から目的信号 $Y_1(l) (l, \tau)$ を読み出し、それに短時間逆フーリエ変換などを施して時間領域の目的信号 $y(l)$ を生成し、メモリ100の記憶領域114に格納する（ステップS64）。

## 【0072】

### （実験結果）

本発明の効果を示すために、マイクロホンに近い主要音を強調する実験を行った。この実験では、図15（a）に示す実験条件でインパルス応答 $h_{qk}(r)$ を測定した。ここでは、カクテルパーティの状況を再現するようにスピーカを配置した。また、全てのスピーカの音量はほぼ等しくなるようにし、特別なスピーカの音量のみが著しく大きくならないようにした。マイクロホンでの混合は、8kHzでサンプリングした6秒間の英語音声、測定したインパルス応答と畳み込むことにより作成した。マイクロホン配置は、図15（a）に示すように3次元的にしたが、本発明を実装したシステム（装置）には、基準マイクロホン（Mic.2）とその他のマイクロホン間の最大距離（3.5cm）のみを情報として与え、詳細なマイクロホン配置は与えていない。各実験において、マイクロホンに近い4つのスピ

一カ位置 (a120, b120, c120, c170) の中から 1 つを目的音源として選び、残りの 3 つのスピーカは無音状態にした。マイクロホンから離れた 6 個のスピーカは、各実験で常に妨害音として作動させた。抽出結果は Input SIR-Output SIR である SIR (signal-to-interference ratio) の改善量によって評価した。この値が大きいほど、より良く目的音声を抽出し、他の妨害音をより抑圧できたことを意味する。これら 2 種類の SIR は、

【数 23】

$$\text{InputSIR} = 10 \log_{10} \frac{\langle |\sum_r h_{11}(r) \cdot s_1(t-r)|^2 \rangle_t}{\langle |\sum_{k \neq 1} \sum_r h_{1k}(r) \cdot s_k(t-r)|^2 \rangle_t} \quad (\text{dB})$$

$$\text{OutputSIR} = 10 \log_{10} \frac{\langle |\sum_r u_{11}(r) \cdot s_1(t-r)|^2 \rangle_t}{\langle |\sum_{k \neq 1} \sum_r u_{1k}(r) \cdot s_k(t-r)|^2 \rangle_t} \quad (\text{dB})$$

によって定義される。ここで、

$$u_{1k}(r) = \sum_{q=1}^M \sum_{\tau=0}^{L-1} w_{1q}(\tau) \cdot h_{qk}(r-\tau)$$

は  $s_k(t)$  から  $y_1(t)$  へのインパルス応答である。

【0073】

実験は、各目的音源位置に対して、7 つの音声 (目的音声 1 個と妨害音声 6 個) からなる組合せを 16 種類作成して行った。図 15 (b) は、ICA のみの場合 (第 2 の実施の形態) と、ICA と時間周波数マスキングの両方を組み合わせた場合 (第 1 の実施の形態) の SIR の平均改善量を示す表である。SIR の改善量は、目的音源の位置によって多少左右されるが、おおむね良好な結果を得ている。位置 a120 と b120 での結果が良いのは、妨害音が異なった位置から到来するからである。位置 c120 と c170 は、2 次元的に見ると、同様の方向から多くの妨害音が到来するため困難な位置に思えるが、実際に位置 c170 での結果は大変良かった。これは、c170 の高さが妨害音と異なり、3 次元状に配列されたマイクロホンによって、本システムが高さの違いを自動的に利用して抽出処理を行っているからである。また、図 15 (b) の表によると、時間周波数マスキングにより性能が改善されることが分かる。時間周波数マスクを決定する式 (26) に対して図 12 (a) に示す 3 種類のパラメータを用いた。より小さい  $\theta_T$  を用いることで、より大きい SIR 改善量が達成されている。しかしながら、小さい  $\theta_T$  によるいくつかの音は、不自然な音 (musical noise) を伴った。実験の結果、多くの場合においてパラメータ  $(\theta_T, g) = (0.333\pi, 20)$  が、十分に妨害音の抑圧を行いながら、かつ自然な音を出力することを確認した。

【0074】

(変形例等)

なお、本発明は上述の各実施の形態に限定されるものではない。例えば、上述した各実施の形態では、信号分離部 130 において M 行 M 列の分離行列  $W(i)$  を算出することとしたが、N 行 M 列など正方行列以外の分離行列  $W(i)$  を算出する構成としてもよい。この場合、基底ベクトルは、分離行列  $W(i)$  の一般化逆行列  $W^+(i)$  (例えばムーア・ペンローズ (Moore-Penrose) 型一般化逆行列) の各列となる。

また、第 1 の実施の形態では、時間周波数マスクを用いて目的信号  $Y_1(i)(l, \tau)$  から妨害信号成分をさらに抑制したマスク処理目的信号  $Y_1(i)'(l, \tau)$  を生成することとしたが、他の方法によって妨害信号成分を抑制したマスク処理目的信号  $Y_1(i)'(l, \tau)$  を生成することとしてもよい。例えば、信号源が 2 個だけである場合、抽出された分離信号  $Y_1(i, \tau)$ 、 $Y_2(i, \tau)$  の大きさを比較し、 $|Y_1(i, \tau)| > |Y_2(i, \tau)|$  ならば  $Y_1(i, \tau)$  をマスク処理目的信号  $Y_1(i)'(l, \tau)$  として抽出し、 $|Y_1(i, \tau)| < |Y_2(i, \tau)|$  ならば  $Y_2(i, \tau)$  をマスク処理目的信号  $Y_1(i)'(l, \tau)$  として抽出する時間周波数マスクを生成し、分離信号  $Y_1(i, \tau)$ 、 $Y_2(i, \tau)$  を要素とするベクトルに掛け合わせることもしてもよい。

【0075】

さらに、上述の各実施の形態では、信号分離部130が独立成分分析(ICA)を用いて分離行列と分離信号とを算出することとしたが、時間周波数マスク(時間周波数毎のマスクを意味し、例えば、1か0の値をとるバイナリマスク)によって観測信号から分離信号を抽出し(例えば、「O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," IEEE Trans. on SP, vol. 52, no. 7, pp. 1830-1847, 2004.」参照)、その算出結果から分離行列を生成することとしてもよい。

また、上述の実施の形態では、ノルム正規化部142bにおいてノルムが1になるような正規化を行ったが、ノルムが1以外の規定値となるように正規化を行ってもよい。

#### 【0076】

さらに、上述の各実施の形態では、フーリエ変換及び逆フーリエ変換によって周波数領域と時間領域との間の変換を行うこととしたが、wavelet変換、DFTフィルタバンク、ポリフェイズフィルタバンクなどを用い、この変換を行うこととしてもよい(例えば、「R. E. Crochiere, L. R. Rabiner, "Multirate Digital Signal Processing," Eaglewood Cliffs, NJ: Prentice-Hall, 1983 (ISBN 0-13-605162-6) )。

また、上述の各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。その他、本発明の趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

また、上述の構成をコンピュータによって実現する場合、各装置が有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、上記処理機能がコンピュータ上で実現される。

#### 【0077】

この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよいが、具体的には、例えば、磁気記録装置として、ハードディスク装置、フレキシブルディスク、磁気テープ等を、光ディスクとして、DVD(Digital Versatile Disc)、DVD-RAM(Random Access Memory)、CD-ROM(Compact Disc Read Only Memory)、CD-R(Recordable)／RW(ReWritable)等を、光磁気記録媒体として、MO(Magneto-Optical disc)等を、半導体メモリとしてEEPROM(Electronically Erasable and Programmable-Read Only Memory)等を用いることができる。

#### 【0078】

また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

#### 【0079】

また、上述した実施形態とは別の実行形態として、コンピュータが可搬型記録媒体から直接このプログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP(Application Service Provider)型のサービスによって、上述の処理を実行する構成としてもよい。なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの(コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等)を含むものとする。

#### 【0080】

また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、本装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実



現することとしてもよい。

【産業上の利用可能性】

【0081】

本発明により、様々な妨害信号が発生する実環境において、目的の信号を精度良く取り出すことが可能となる。音信号に対する応用例としては、音声認識器のフロントエンドとして働く目的信号抽出システムなどが挙げられる。話者とマイクロホンとが離れた位置にあり、マイクロホンが話者の音声以外の音を集音してしまうような状況でも、そのようなシステムを使うことで、話者の音声のみを抽出して正しく音声を認識することができる。

【図面の簡単な説明】

【0082】

【図1】第1の実施の形態における信号抽出装置のハードウェア構成を例示したブロック図。

【図2】第1の実施の形態における信号抽出装置のブロック図。

【図3】(a)は図2における目的信号選択部の詳細構成を例示したブロック図。(b)は(a)における基底ベクトルクラスタリング部の詳細構成を例示したブロック図。

【図4】(a)は図2における時間周波数マスキング部の詳細構成を例示したブロック図。(b)は(a)におけるマスク生成部の詳細を例示したブロック図。

【図5】第1の実施の形態における信号抽出処理の全体を説明するためのフローチャート。

【図6】第1の実施の形態における目的信号選択部の処理の詳細を説明するためのフローチャート。

【図7】(a)はステップS12の正規化処理の詳細を説明するためのフローチャート。(b)はステップS13の正規化処理の詳細を説明するためのフローチャート。

【図8】目的信号の選択手順の詳細を説明するためのフローチャート。

【図9】パラメータd毎の正規化基底ベクトルの要素 $A_{qp}''(l)$ とその偏角 $\arg[A_{qp}''(l)]$ との関係を説明するための複素平面図。

【図10】パラメータd毎の正規化基底ベクトルの要素 $A_{qp}''(l)$ とその偏角 $\arg[A_{qp}''(l)]$ との関係を説明するための複素平面図。

【図11】図5におけるステップS4の詳細を説明するためのフローチャート。

【図12】(a)は時間周波数マスク $M(l, \tau)$ の例示。(b)はある時間周波数位置 $(l, \tau)$ における白色化空間を説明するための概念図。

【図13】第2の実施の形態における信号抽出装置のブロック図。

【図14】第2の実施の形態における信号抽出装置の処理を説明するためのフローチャート。

【図15】(a)は実験条件を示す図。(b)は実験結果を示した表。

【符号の説明】

【0083】

1, 200 信号抽出装置

【書類名】 図面

【図 1】

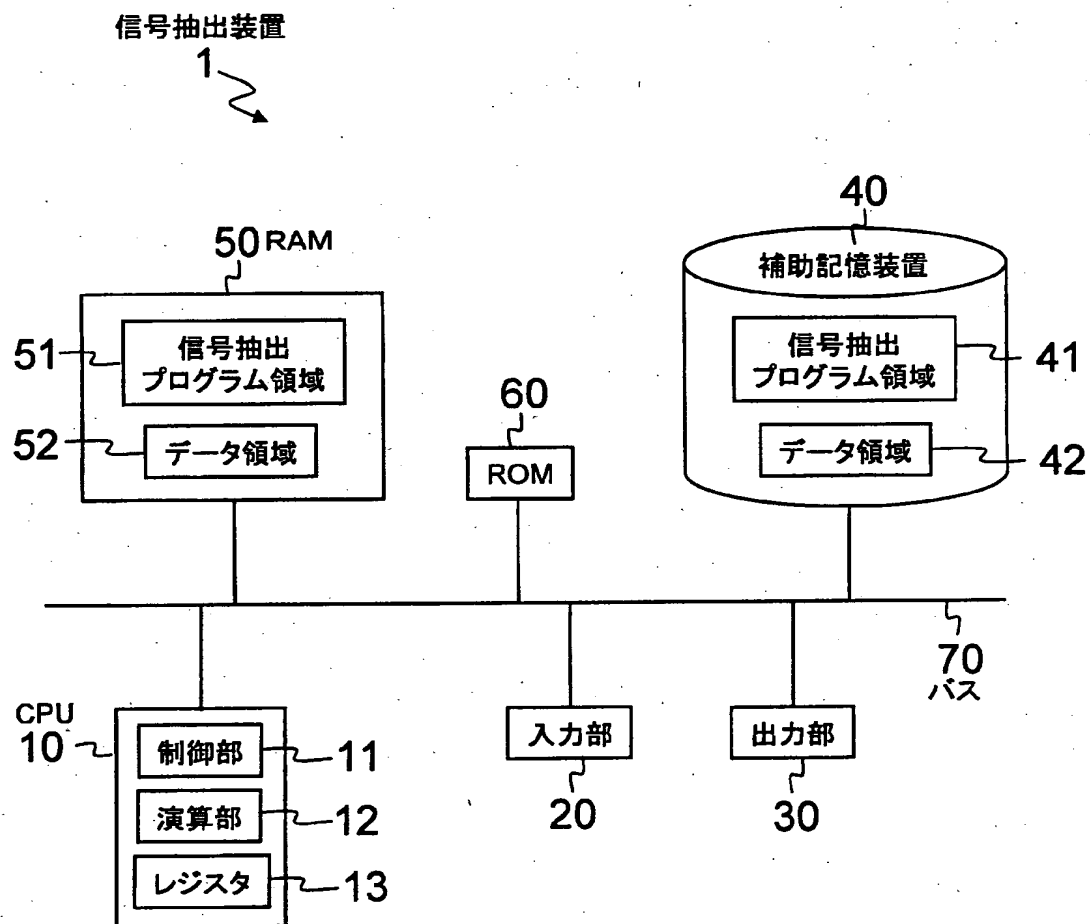
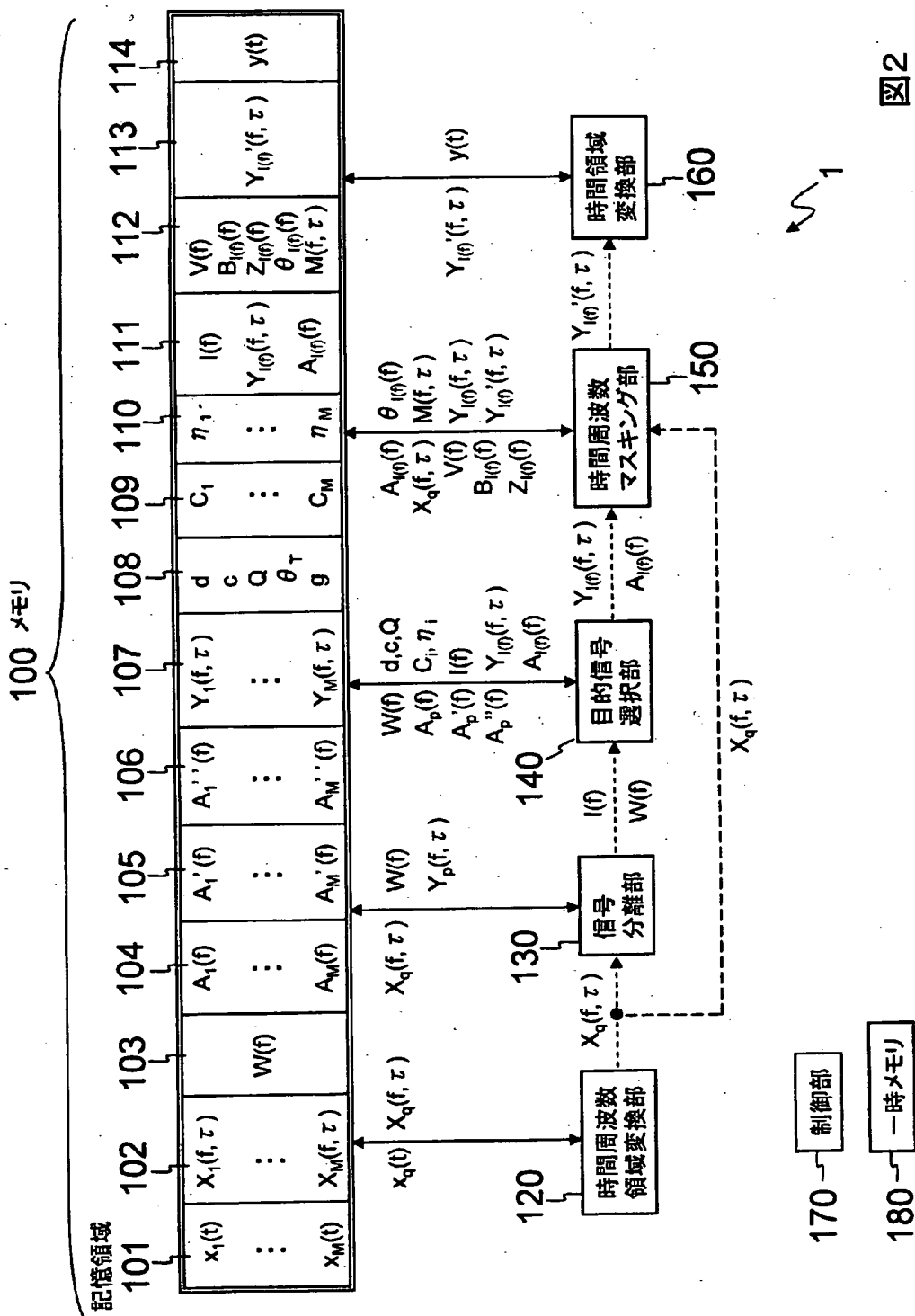
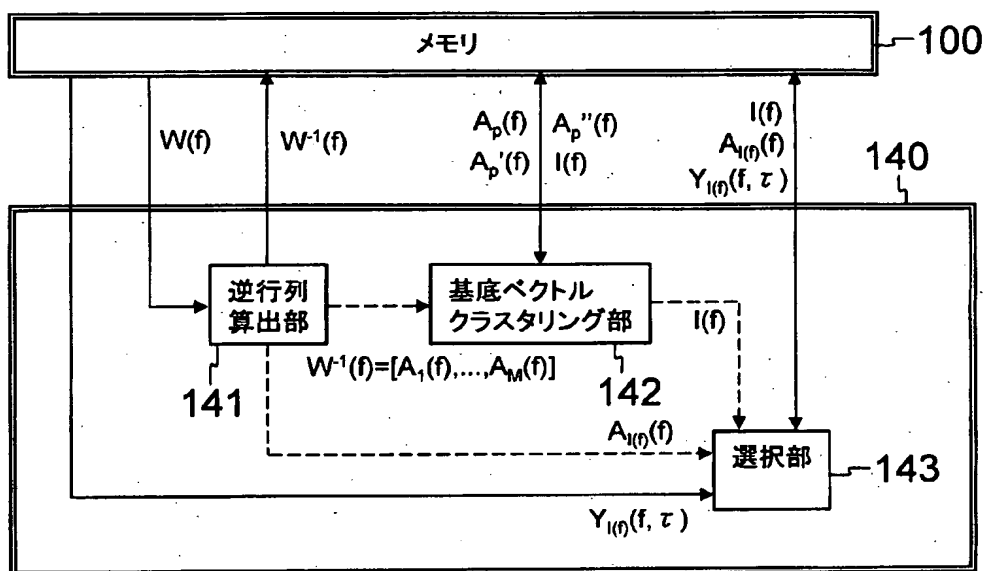


図1



【図 3】

(a)



(b)

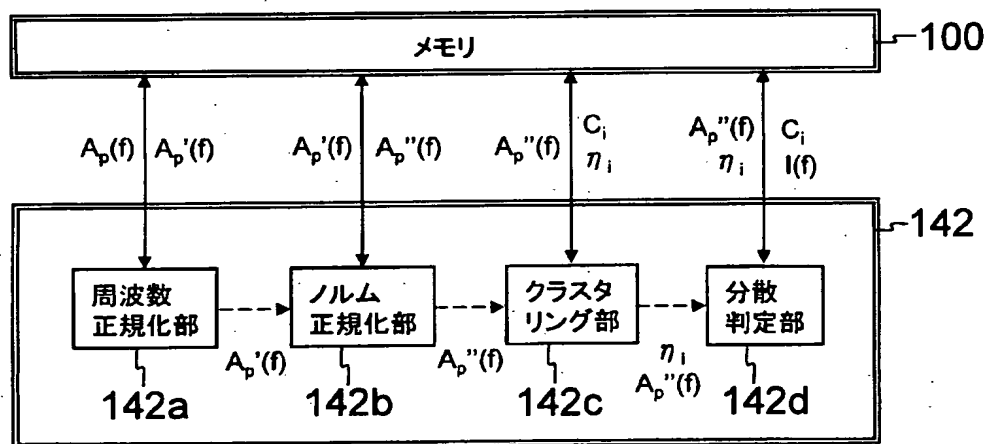
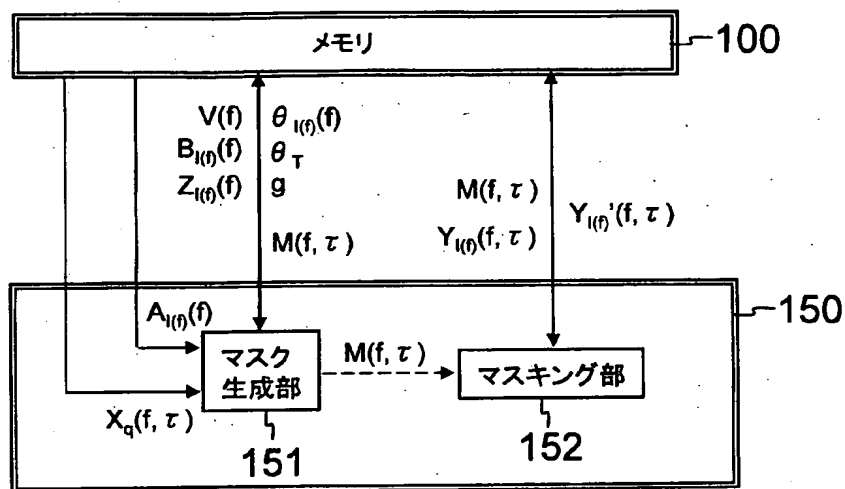


図3

【図 4】

(a)



(b)

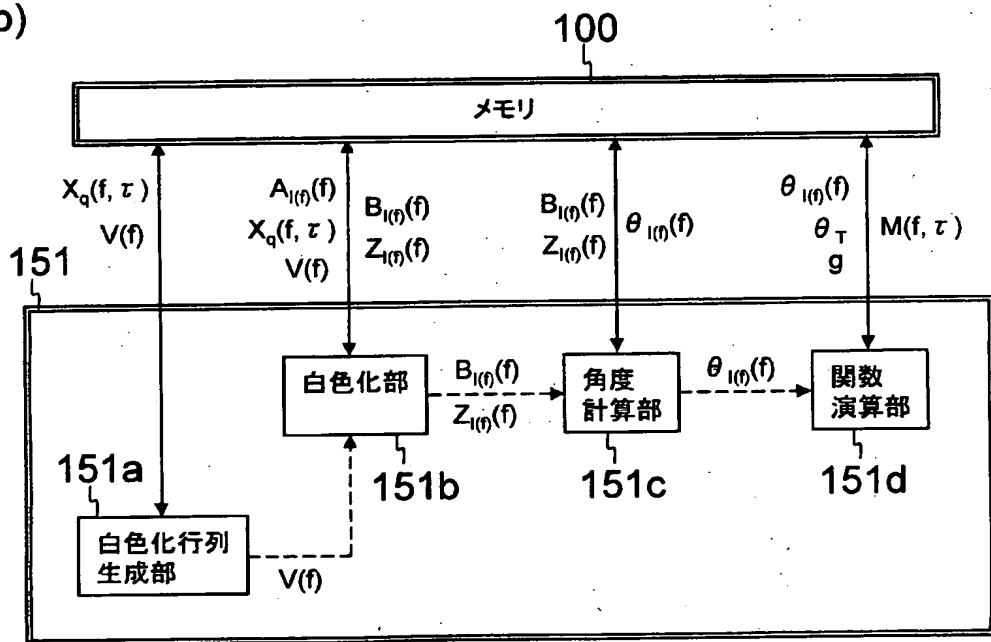


図4

【図 5】

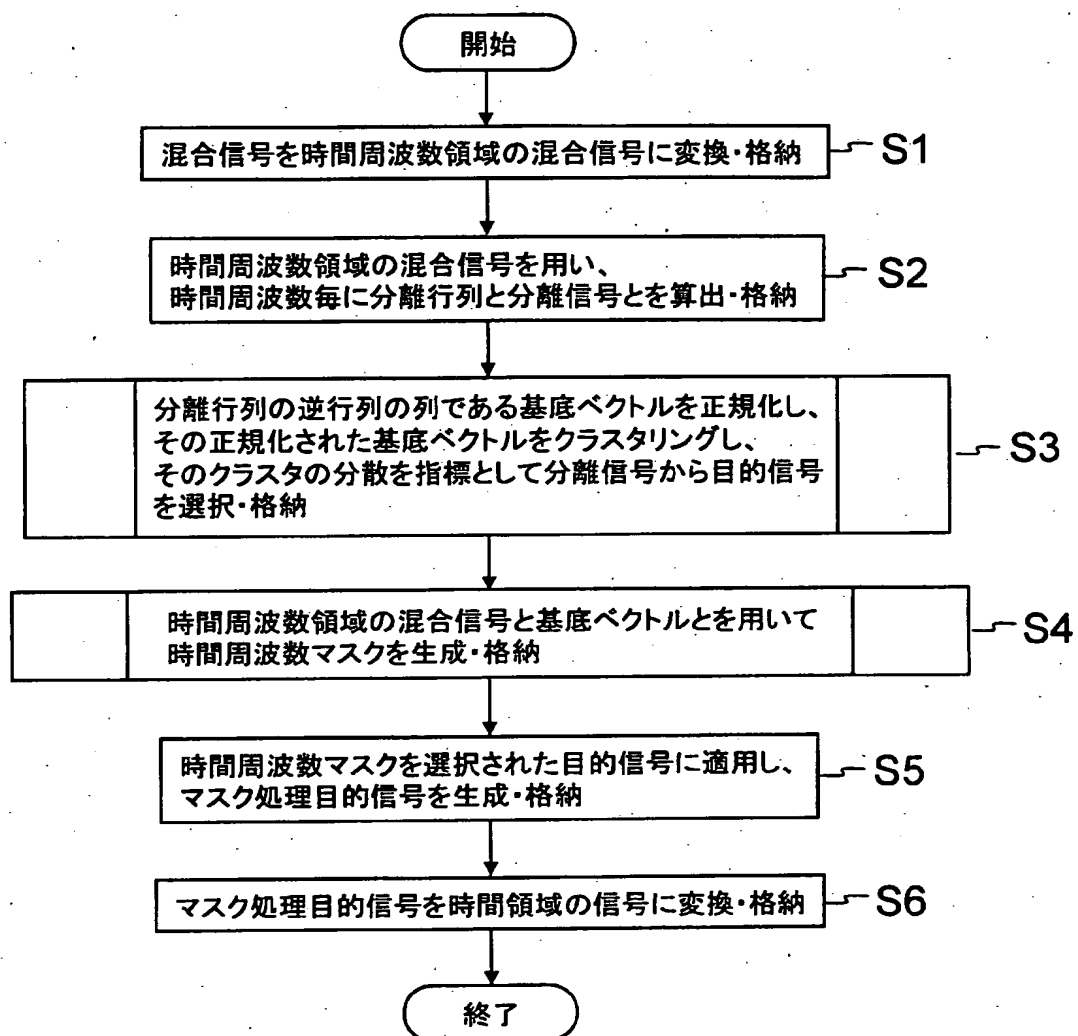


図5

【図 6】

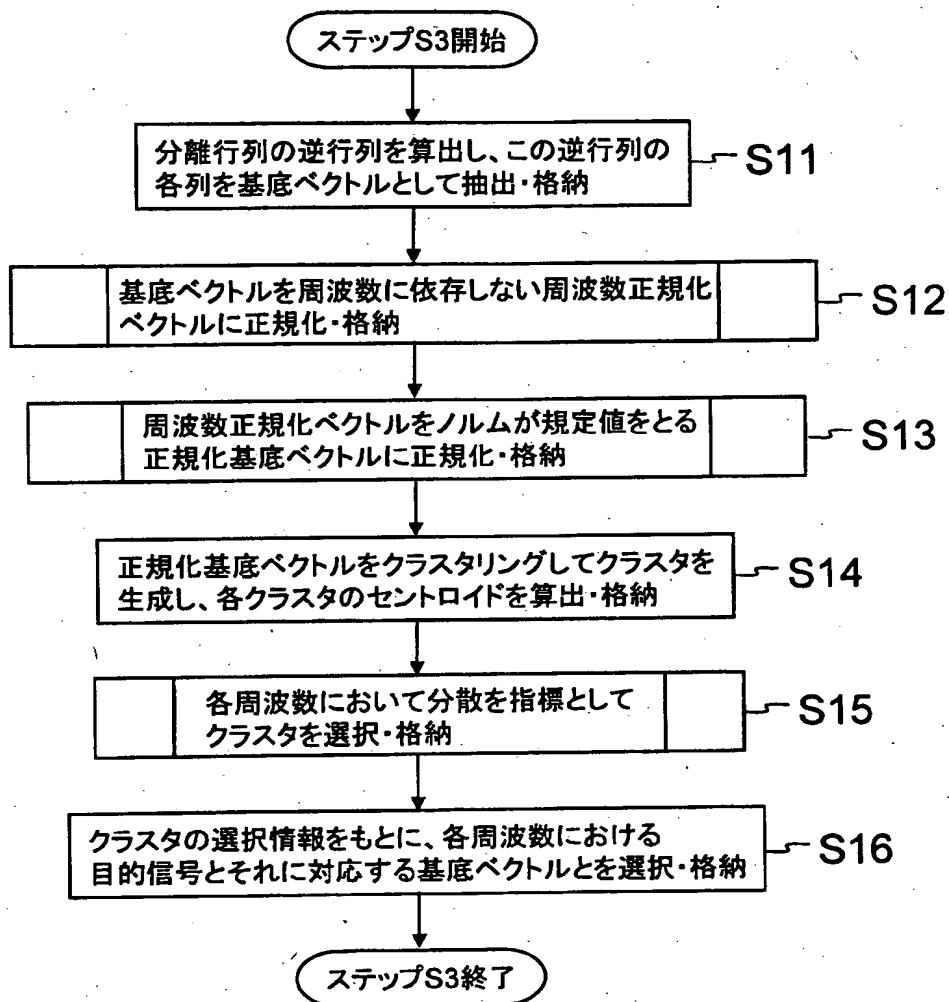
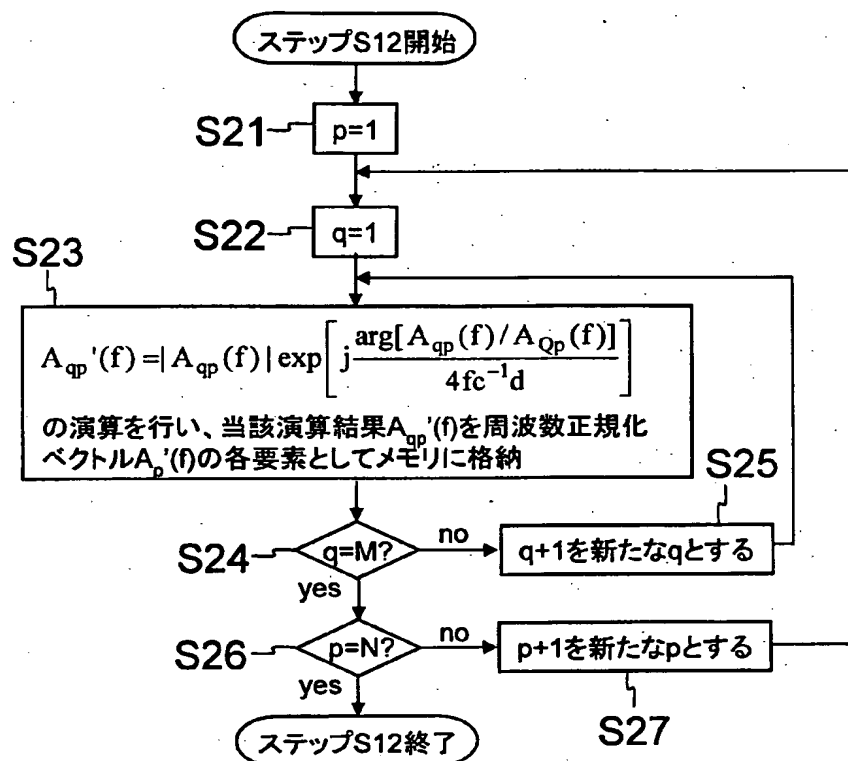


図6

【図 7】

(a)



(b)

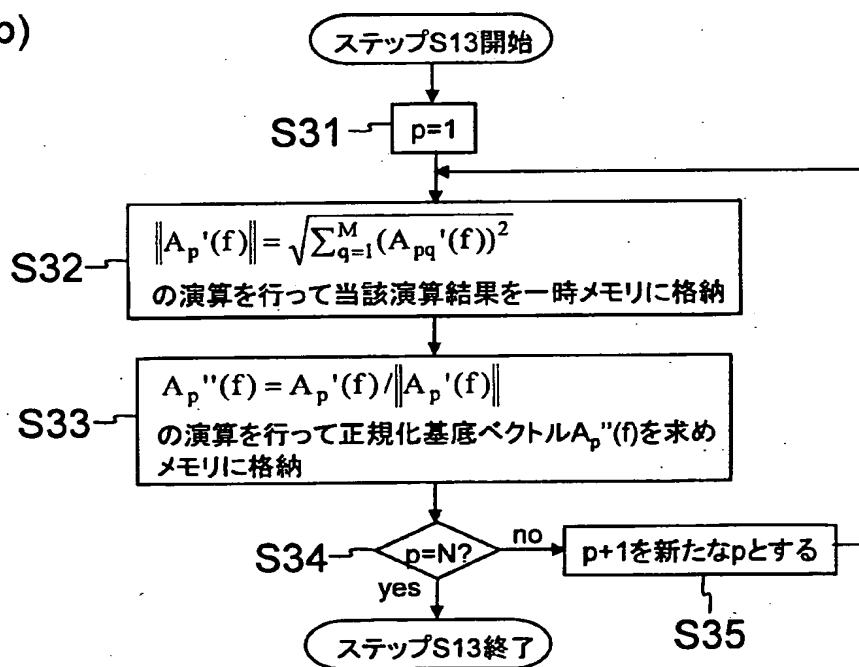


図7



【図 8】

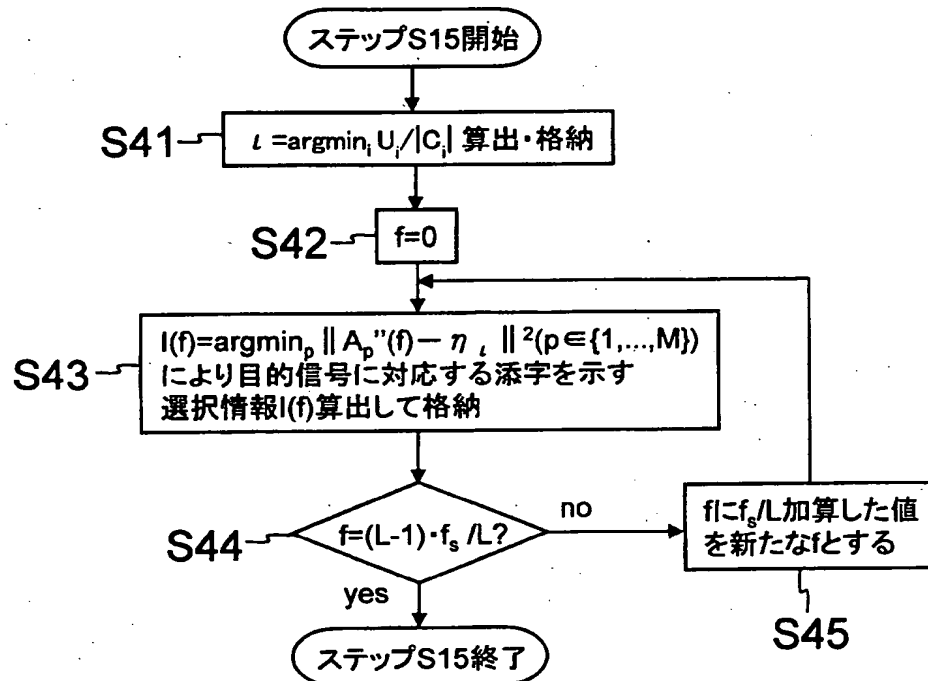
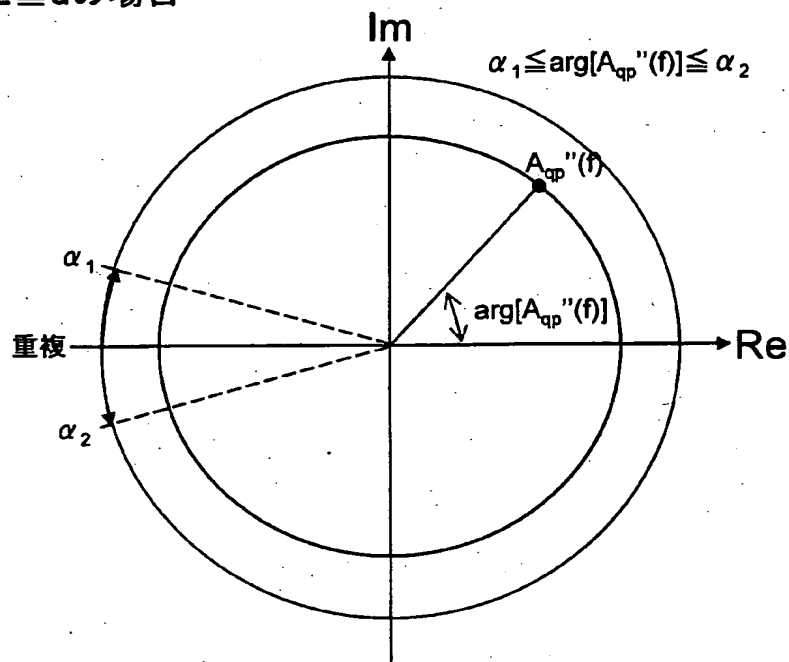
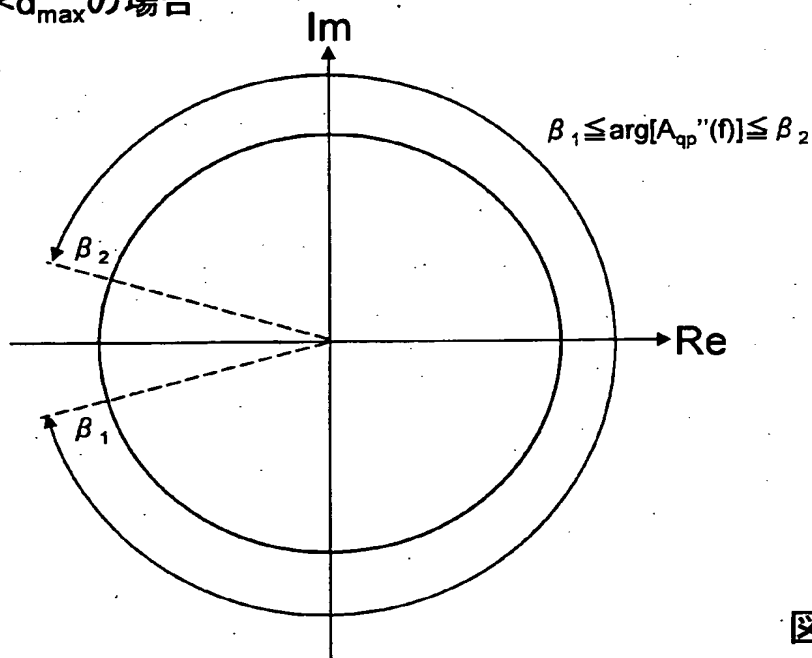


図 8

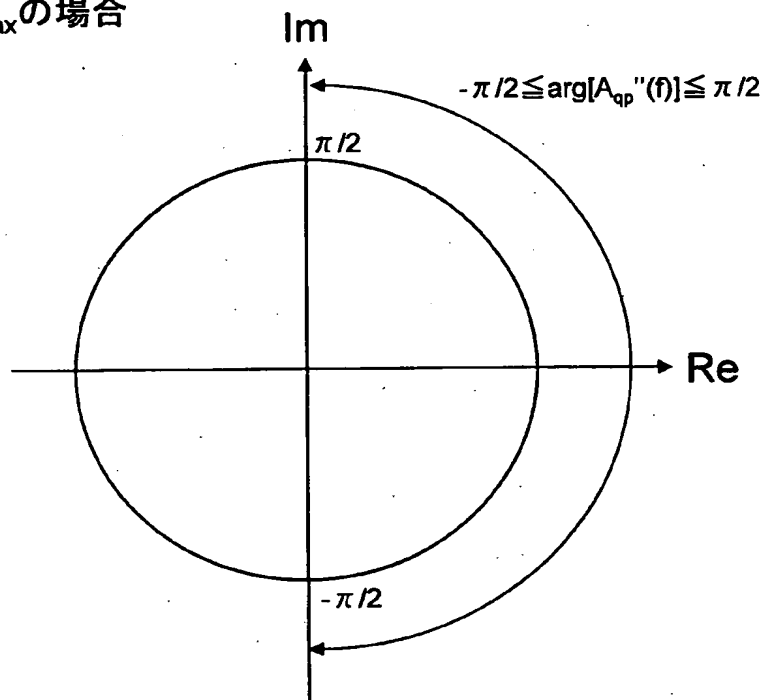
(a)  $d_{\max}/2 \geq d$  の場合



(b)  $d_{\max}/2 < d < d_{\max}$  の場合



(a)  $d=d_{\max}$  の場合



(b)  $d > d_{\max}$  の場合

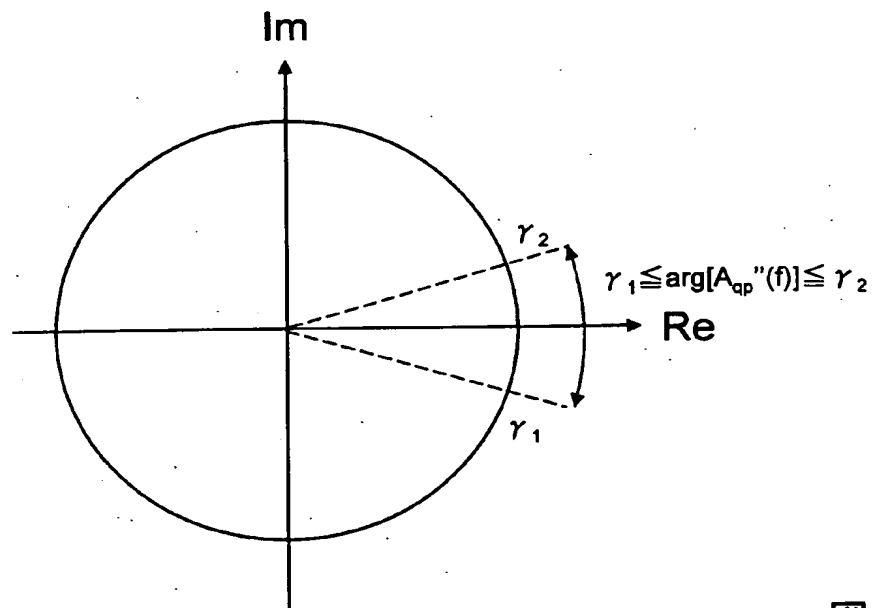
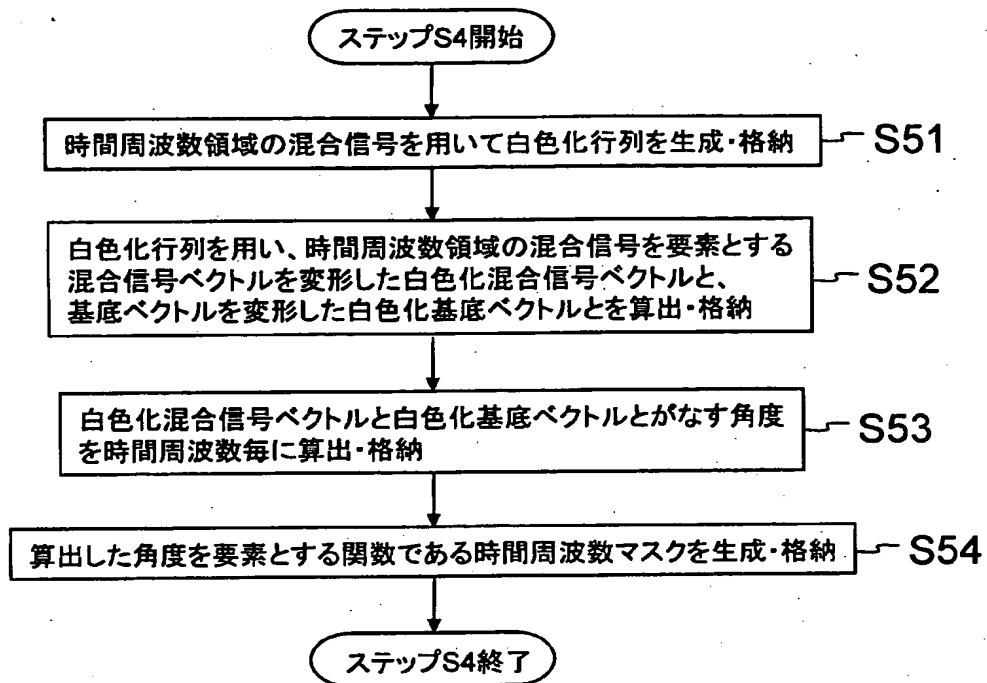
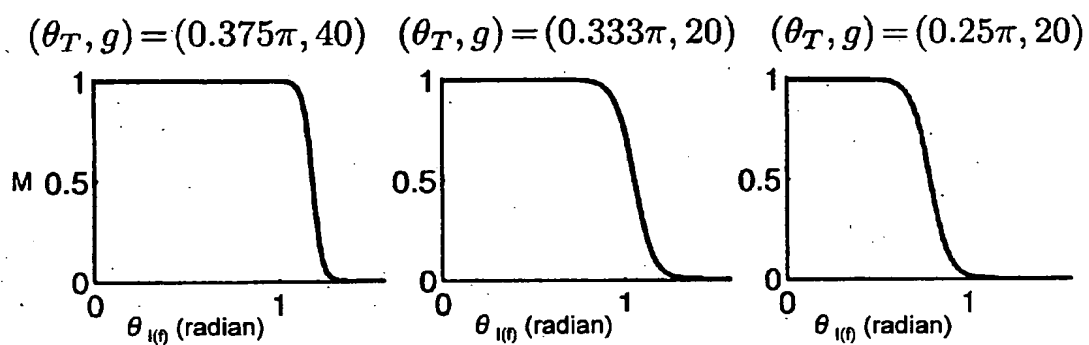


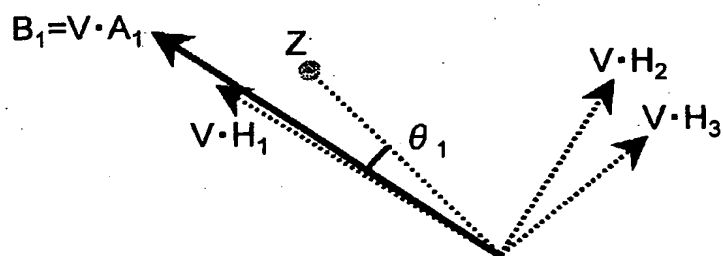
図10



(a)



(b)



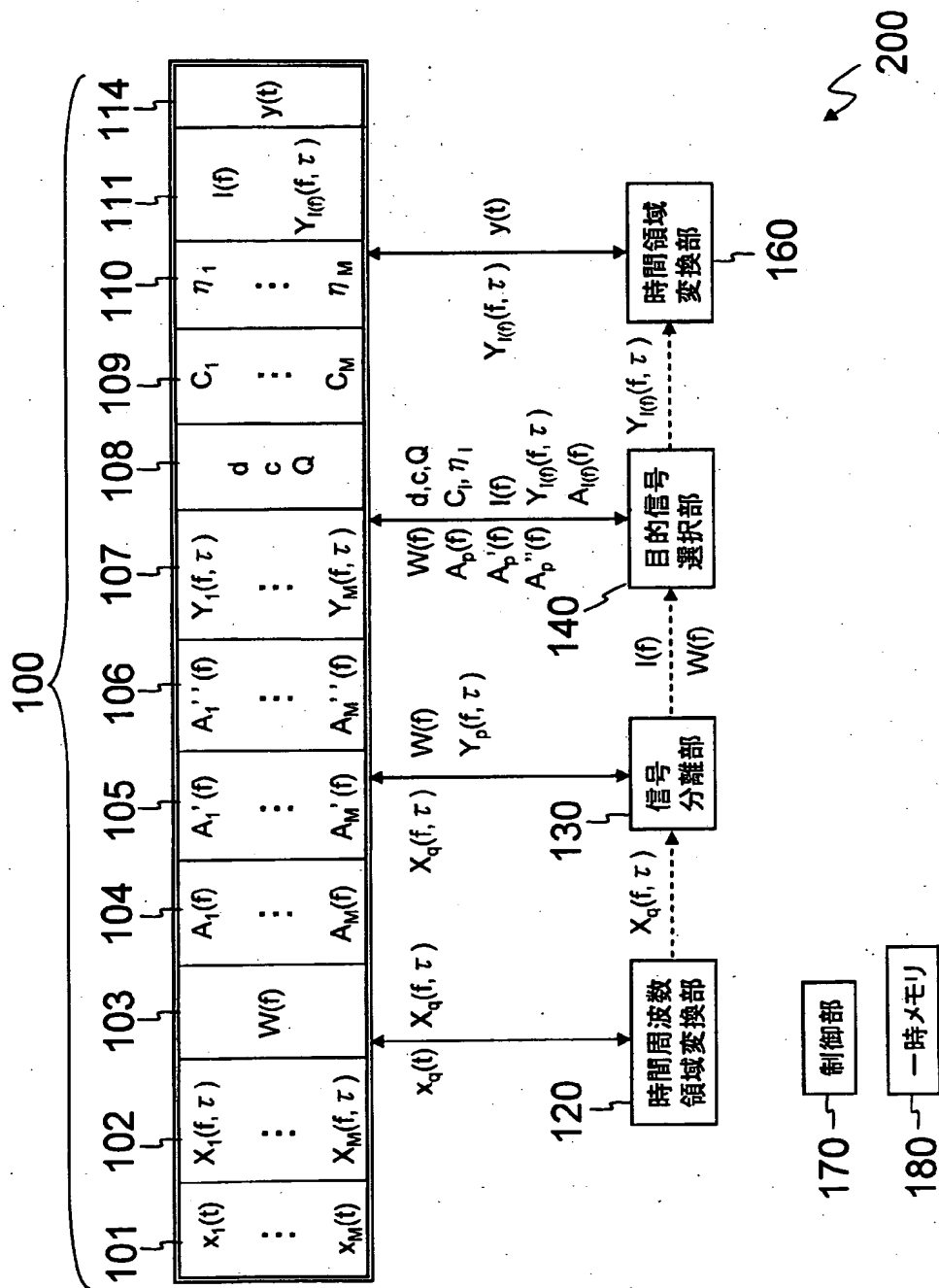


図13

【図 14】

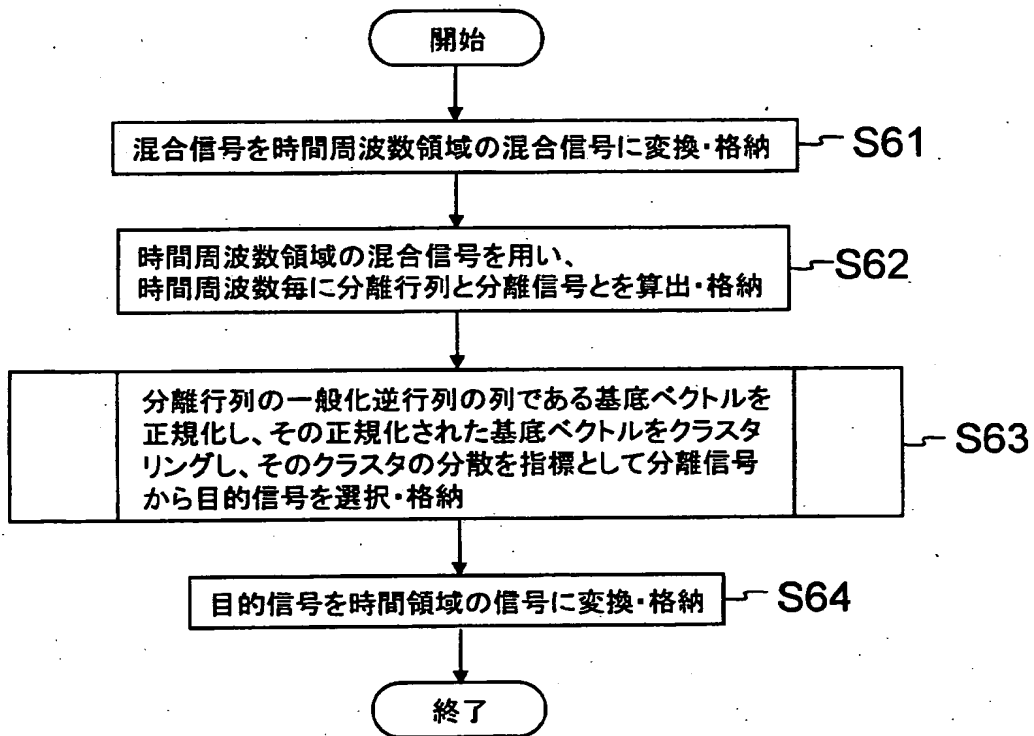
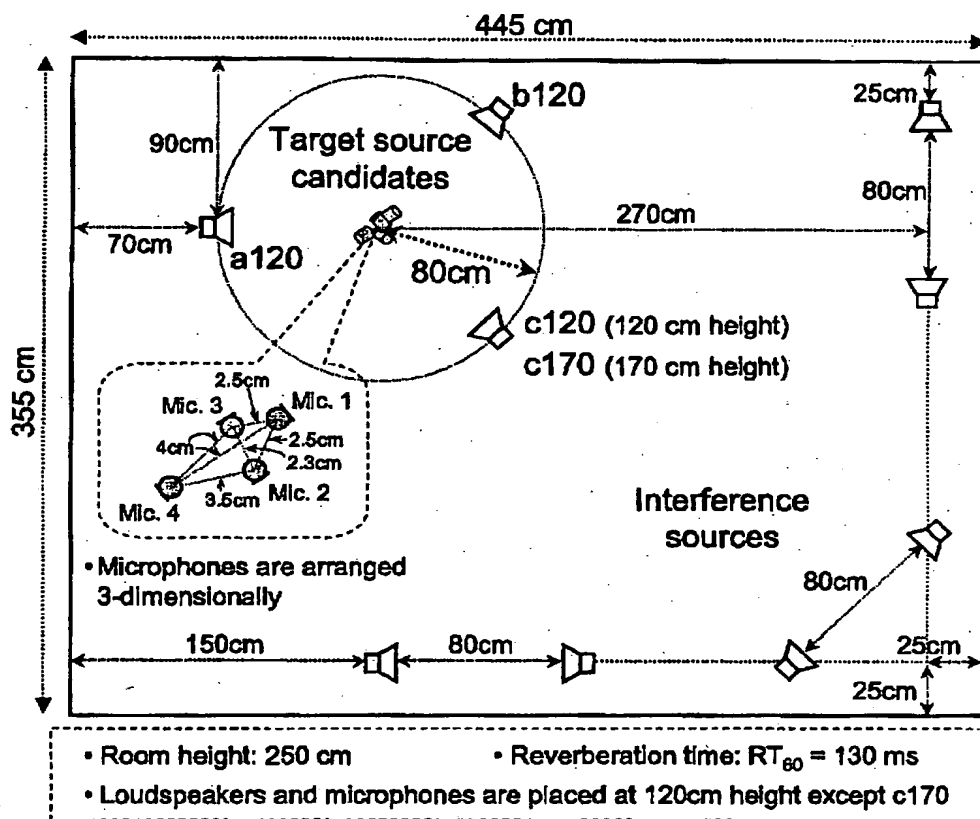


図14

(a)



(b)

目的音声位置	a120	b120	c120	c170
ICA のみ	11.7 dB	11.7 dB	9.0 dB	12.9 dB
マスキングを併用 ( $0.375\pi, 40$ )	15.5 dB	14.3 dB	12.4 dB	16.7 dB
マスキングを併用 ( $0.333\pi, 20$ )	16.9 dB	15.4 dB	14.0 dB	18.1 dB
マスキングを併用 ( $0.25\pi, 20$ )	19.7 dB	17.9 dB	16.9 dB	20.8 dB



【書類名】 要約書

【要約】

【課題】 目的信号の方向に関する情報を必要とせずに、複数の信号源から発せられた信号が混合された混合信号から目的信号を抽出する。

【解決手段】 複数の信号源から発せられた信号が混合された混合信号  $x_q(i)$  を時間周波数領域の混合信号  $X_q(i, \tau)$  に変換し、時間周波数領域の混合信号  $X_q(i, \tau)$  を用い、周波数毎に分離行列  $W(i)$  と分離信号  $Y_p(i, \tau)$  とを算出する。次に、分離行列  $W(i)$  の一般化逆行列の列である基底ベクトル  $A_p(i)$  を正規化し、その正規化された基底ベクトル  $A_p^{(1)}(i)$  をクラスタリングし、そのクラスタの分散を指標として分離信号  $Y_p(i, \tau)$  から目的信号  $Y_1(i, \tau)$  を選択する。そして、好ましくは、時間周波数マスク  $M(i, \tau)$  を生成し、この時間周波数マスク  $M(i, \tau)$  を選択された目的信号  $Y_1(i, \tau)$  に適用し、マスク処理目的信号  $Y_1(i, \tau)$  を生成する。

【選択図】 図2

出願人履歴

000004226

19990715

住所変更

591029286

東京都千代田区大手町二丁目3番1号  
日本電信電話株式会社